

UNIVERSITY OF TORONTO SCARBOROUGH
Department of Computer and Mathematical Sciences
DECEMBER 2013 EXAMINATIONS

STAB22H3 Statistics I

Duration: 3 hours

Last Name: _____ First Name: _____

Student number: _____

Aids allowed:

- Two handwritten letter-sized sheets (both sides) of notes prepared by you
- Non-programmable, non-communicating calculator

Standard Normal, t and the binomial distribution tables are attached at the end.

This test is based on multiple-choice questions. There are 44 questions. All questions carry equal weight. On the Scantron answer sheet, ensure that you enter your last name, first name (as much of it as fits), and student number (in “Identification”).

Mark in each case the best answer out of the alternatives given (**which means the numerically closest answer if the answer is a number and the answer you obtained is not given.**)

Also before you begin, complete the signature sheet, but sign it only when the invigilator collects it. The signature sheet shows that you were present at the exam.

Good luck!!

ExamVersion:

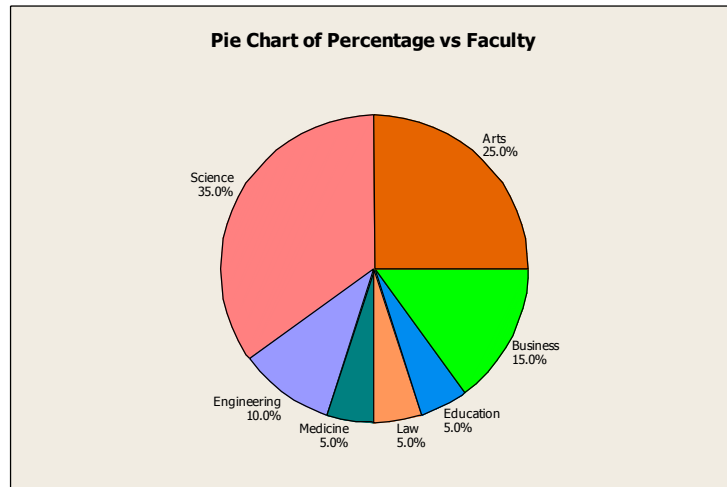
1. In a study of acupuncture for treating pain, 100 volunteers were recruited. Half were randomly assigned to receive acupuncture and the other half to receive a sham acupuncture treatment. The patients were followed for 6 months and the treating physician measured their degree of pain relief. The patients did not know which treatment they actually received, but the treating physicians were aware of who was getting acupuncture and who wasn't. Which of the following best describes this study?
 - (a) This is a retrospective observational study.
 - (b) This is a randomized block design.
 - (c) This is a randomized single-blind experiment.
 - (d) This is a randomized double-blind experiment.
 - (e) This is a prospective observational study.

2. Which of the following pairs of events (denoted by A and B) are disjoint and independent?
 - (a) Toss two coins, A = first toss is head, B = second toss is head
 - (b) Roll a die, A = observe an odd number, B = observe an even number
 - (c) Take two cards one by one from a deck of 52 cards, A = first card is red, B = both cards are red
 - (d) There are three kids in a family, A = two kids are girls, B = two kids are twins
 - (e) None of the above

3. In a group of 100 people, 40 own a cat, 25 own a dog, and 15 own a cat and a dog. Find the probability that a person chosen at random, owns a dog, given that he owns a cat.
 - (a) 0.375
 - (b) 0.6
 - (c) 0.25
 - (d) 0.4
 - (e) 0.15

4. Two stores (store A and store B) sell watermelons. The weights of watermelons at store A has a Normal distribution with mean 21 pounds and standard deviation 3 pounds. The weights of watermelons at store B has a Normal distribution with mean 18 pounds and standard deviation 2 pounds. The weights of watermelons at the two stores are independent. If you select a melon at random at each store, what is the probability that melon you selected at store B will be heavier? Choose the closest.
- (a) 0.1
 - (b) 0.2
 - (c) 0.3
 - (d) 0.4
 - (e) 0.5
5. The regression equation relating midterm score (x) and final exam score (y) in a course is $y = 20 + 0.8x$. Which of the following is a valid conclusion based on this information?
- (a) If Adam has scored 10 points higher than Bob on the midterm, then Adam will score 8 points higher than Bob on the final exam.
 - (b) Students who score 80 on the midterm are predicted to score 84 on the final.
 - (c) If a student scores 0 on the midterm, then that student will score 20 on the final.
 - (d) On average, the students' scores on the final exam will only be 80% of what they were on the midterm.
 - (e) This equation shows that the midterm and the final exam scores are not related.
6. A test of significance with $\alpha = 0.05$ results in a P-value of 0.1. What do you conclude?
- (a) Reject the null hypothesis.
 - (b) Reject the alternative hypothesis.
 - (c) Do not reject the alternative hypothesis.
 - (d) Do not reject the null hypothesis.
 - (e) Accept the null hypothesis.

7. The pie chart below shows the percentage of students in each faculty at a university.



If the number of students in the faculty of Arts is 3000, then how MANY students are there in the faculty of Science?

- (a) 3400
 - (b) 3600
 - (c) 3800
 - (d) 4000
 - (e) 4200
8. The two-way table below classifies the members of a fitness club by gender and food habits. Use this information to answer this question and the next question.

	Women	Men
Vegetarian	9	3
Non-vegetarian	8	10

If a member is selected at random from the club, what is the probability that this person is a vegetarian female?

- (a) $\frac{3}{4}$
- (b) $\frac{3}{10}$
- (c) $\frac{1}{10}$
- (d) $\frac{4}{9}$
- (e) $\frac{9}{17}$

9. Using the information in question 8 above, what is the probability that a randomly selected member is non-vegetarian?

- (a) 2/5
- (b) 1/3
- (c) 4/15
- (d) 3/5
- (e) 8/17

10. The regression equation and the summary statistics given below were obtained from a regression analysis of the relation between murder rate (y , the number of murders per 1,000,000 inhabitants per annum) and unemployment(x , the percentage of the labour force unemployed) using data for 20 cities.

Descriptive Statistics: x , y

Variable	N	Mean	StDev	Minimum	Q1	Q3	Maximum
x	20	deleted	1.207	4.900	6.050	8.125	9.300
y	20	20.57	9.88	5.30	12.88	26.63	40.70

Regression Analysis: y versus x

The regression equation is $y = - 28.5 + 7.08 x$

What is the correlation between x and y ?

- (a) 0.035
- (b) 0.0708
- (c) 0.122
- (d) 0.865
- (e) 0.988

11. Using the information in question 10 above, calculate the average unemployment rate (i.e. the value of \bar{x}) of these 20 cities. Note: This value has been deleted in the summary statistics given in question 10 above but can be calculated from other information given.

Which of the following numbers is closest to this value?

- (a) 7
- (b) 14
- (c) 21
- (d) 28
- (e) 49

12. One of the cities in the data set used in question 10 above, had unemployment rate 6.0 (i.e. $x = 6.0$) and murder rate 14.5. What is the residual for this city?
- (a) 0.52
 - (b) -68.16
 - (c) -27.98
 - (d) 43
 - (e) -96.66
13. Which of the following procedures best describes a method to select a simple random sample?
- (a) Randomly select half of the sample from females and the remaining half from males.
 - (b) Select an individual from every fourth house on a street.
 - (c) Assign each individual in the population a unique number and use a computer or a random number table to randomly generate numbers for selection.
 - (d) Select every individual with a surname beginning with the letter S.
 - (e) Select every 20th individual from a list of patients registered with a GP.
14. X and Y are two independent random variables and $Var(X) = 0.61$ and $Var(Y) = 0.21$. Find $SD(X - Y)$ (i.e. σ_{X-Y}).
- (a) 0.6451
 - (b) 0.6324
 - (c) 0.9055
 - (d) 0.5727
 - (e) 0.3227
15. In a large mall a survey was taken. It was found that in a random sample of 45 women over the age of 25, 15 had children. Find the 90% confidence interval for the population proportion of women over the age of 25 in the mall who have children. (Choose the closest answer.)
- (a) (0.0208, 0.0653)
 - (b) (0.4337, 0.6774)
 - (c) (0.3333, 0.5555)
 - (d) (0.1374, 0.1626)
 - (e) (0.2177, 0.4489)

16. In a regression study, which of the following best describes the square of the correlation r^2 between the two variables?
- (a) The amount of error that is incurred from using a regression model to make predictions
 - (b) The amount of the variability explained by the standard deviations of the response and explanatory variables
 - (c) The proportion of the total variation in the response variable explained by the straight line relationship with the explanatory variable
 - (d) The sum of squares of the residuals
 - (e) None of the above
17. Two independent random variables have the following distributions:

x	-1	0	1
p	0.2	0.3	0.5

y	1	2
p	0.3	0.7

Find the mean of $Y - 2X$ (i.e. find $E(Y - 2X)$).

- (a) 1.1
 - (b) -3.1
 - (c) 0.3
 - (d) 1.7
 - (e) 1.4
18. It is widely thought that there is a high incidence of disability among the homeless population. A random sample of 110 homeless people contained 84 who were disabled on one or more categories (such as psychiatric disability, medical disability etc.). Let p denote the proportion of all homeless people having one or more types of disability. Which of the following numbers is the closest to the value of the test Z-statistic for testing the null hypothesis $H_0 : p = 0.75$ against $H_a : p \neq 0.75$?
- (a) 0.0
 - (b) 0.3
 - (c) 0.6
 - (d) 0.9
 - (e) 1.2

19. The events A and B are such that $P(A) = 2P(B)$, $P(A \text{ and } B) = 0.08$ and A and B are independent. Find $P(A \text{ or } B)$.
- (a) 0.68
 - (b) 0.6
 - (c) 0.4
 - (d) 0.52
 - (e) 0
20. The heights of women in a certain population have a Normal distribution with mean 64.5 inches and standard deviation 2.5 inches. We select three women at random from this population. Assume that their heights are independent. Find the probability that the tallest of these three women will be taller than 67 inches. Which of the following numbers is closest to this probability?
- (a) 0.0
 - (b) 0.2
 - (c) 0.4
 - (d) 0.6
 - (e) 0.8
21. Which of the following values is closest to the interquartile range for the standard normal distribution?
- (a) 0
 - (b) 0.5
 - (c) 1.3
 - (d) 2.5
 - (e) 3

22. The distribution of bachelor's degrees conferred by a local college is listed below, by major.

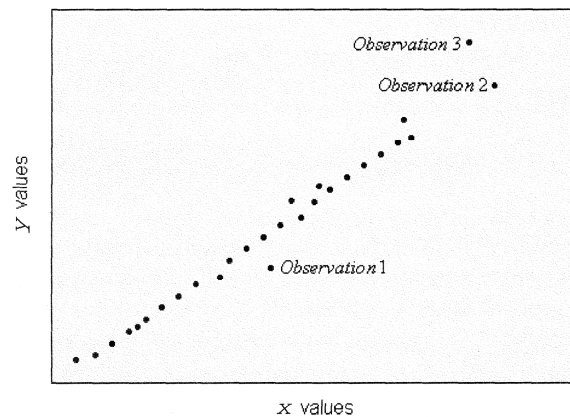
Major	Frequency
English	2,073
Mathematics	2,164
Chemistry	318
Physics	856
Liberal Arts	1,358
Business	1,676
Engineering	868
Total	9,313

- What is the probability that a randomly selected bachelor's degree is not in Mathematics?
- (a) 0.232
 - (b) 0.768
 - (c) 0.303
 - (d) 0.682
 - (e) 0.889
23. You have sampled 25 students to find the mean SAT scores. A 95% confidence interval for the mean SAT score is 900 to 1100. Which of the following statements gives a valid interpretation of this interval?
- (a) 95% of the 25 students have a mean score between 900 and 1100.
 - (b) 95% of the population of all students have a score between 900 and 1100.
 - (c) If this procedure were repeated many times, 95% of the sample means would be between 900 and 1100.
 - (d) If this procedure were repeated many times, 95% of the resulting confidence intervals would contain the true mean SAT score.
 - (e) If 100 samples were taken and a 95% confidence interval was computed, 5 of them would be in the interval from 900 to 1100.

24. A manufacturer of light bulbs claims that their light bulbs have a mean life of 1520 hours with an unknown standard deviation. A random sample of 30 such bulbs is selected for testing. You want to test the null hypothesis that the population mean life is less than the manufacturer claims. If the sample produces a mean value of 1505 hours and a sample standard deviation of 86, what can you conclude? You may assume that all the necessary assumptions for the related test are satisfied by the data.
- (a) We will reject the null hypothesis at $\alpha = 0.01$
 - (b) We will reject the null hypothesis at $\alpha = 0.05$
 - (c) We will reject the null hypothesis at $\alpha = 0.1$
 - (d) We will reject the null hypothesis at $\alpha = 0.001$
 - (e) We fail to reject the null hypothesis at $\alpha = 0.1$
25. The manufacturer of a bag of sweets claims that there is a 90% chance that the bag contains some toffees. If 20 bags are chosen, how many bags with toffees would you expect to have?
- (a) 20
 - (b) 2
 - (c) 9
 - (d) 18
 - (e) None
26. In brief, what does the Central Limit Theorem say?
- (a) The area under a Normal density curve is one.
 - (b) Measures of central tendency should always be computed with and without outliers.
 - (c) For sufficiently large sample size, the sampling distribution of \bar{X} is approximately Normal.
 - (d) Confidence intervals have zero margin of error for large sample sizes.
 - (e) In the long run, the average outcome gets close to the distribution mean.
27. A consumer product agency tests mileage (miles per gallon) for a sample of automobiles using each of four different types of gasoline. Which of the following statements regarding this study is true?
- (a) There are four explanatory variables and one response variable in this study.
 - (b) There is only one explanatory variable with four levels and one response variable in this study.
 - (c) Miles per gallon is the only explanatory variable in this study.
 - (d) This study has a single response variable with four levels.
 - (e) None of the above statements is true.

28. The heights (y) of 50 men and their shoes sizes (x) were obtained. The variable height is measured in centimetres (cm) and the shoe sizes of these men ranged from 8 to 13. From these 50 pairs of observations, the least squares regression line predicting height from shoe size was computed to be $\hat{y} = 130.455 + 4.7498x$. What height would you predict for a man with a shoe size of 13?
- (a) 130.46 cm
 - (b) 192.20 cm
 - (c) 182.70 cm
 - (d) This regression line cannot be used to predict the height of a man with a shoe size of 13
 - (e) None of the above
29. The weekly amount of money spent on maintenance and repairs by a company has a Normal distribution with mean \$400 and standard deviation \$ 20. If \$ 450 is budgeted for next week, what is the probability that the actual cost will exceed the budget? Choose the value closest to this probability.
- (a) 0.002
 - (b) 0.006
 - (c) 0.025
 - (d) 0.01
 - (e) 0.5
30. The exam scores in a class has a Normal distribution. 2.5 percent of the class failed (i.e. scored below 50) and 16 percent of the class got A's. A is the highest grade in this exam and a score of 86 or above qualifies for an A grade. Based on 68-95-99.7 % rule, what is the mean of this Normal distribution?
- (a) 56
 - (b) 62
 - (c) 68
 - (d) 74
 - (e) 78

31. Government data show that 10% of males under age 25 are unemployed. A random sample is taken of 400 males who are in the labor force and under age 25. Find the probability that the sample unemployment rate is 0.12 or more. (Hint: Use Normal approximation)
- (a) 0.1023
 (b) 0.0918
 (c) 0.9082
 (d) 0.5470
 (e) 0.4530
32. The scatterplot of two variables x and y is shown below. This scatterplot shows three observations that appear to be somewhat unusual.



Read the following statements labeled I, II, III and IV, carefully.

- I Observations 1 and 2 are not influential.
 II Observations 2 and 3 have high leverage.
 III Observations 2 and 3 have large residuals.
 IV Out of the three observations, only observation 3 is an outlier.

Which of the above statements are true?

- (a) Only I is true.
 (b) Only I and II are true.
 (c) only I, II, and III are true.
 (d) only IV is true.
 (e) All are true.

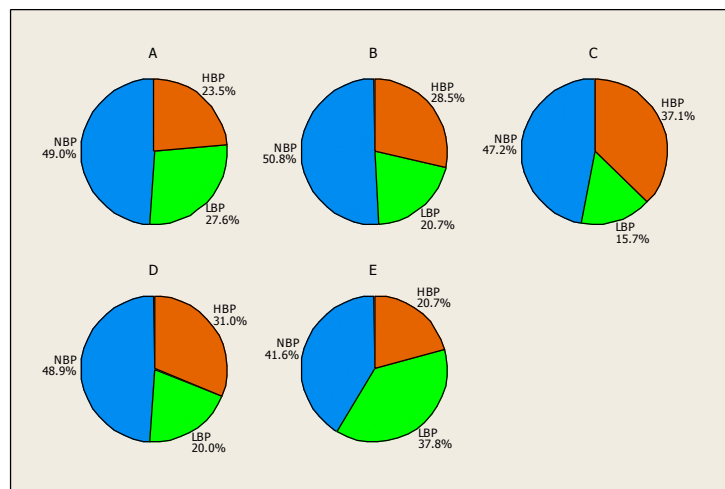
33. A number calculated with complete population data and quantifies a characteristic of the population is called ...
- (a) Datum
 - (b) Statistics
 - (c) Population
 - (d) Stratum
 - (e) Parameter
34. A missile protection system consists of 3 radar sets operating independently, each with probability 0.9 of detecting a missile entering the zone that is covered by the system. If a missile enters the zone, what is the probability that at least one radar set will detect the missile?
- (a) 0.9
 - (b) 0.729
 - (c) 0.271
 - (d) 0.99
 - (e) 0.999
35. In clinical trials a certain drug has a 10% success rate of curing a known disease. If 15 people are known to have the disease. What is the probability of at least 2 being cured?
- (a) 0.5491
 - (b) 0.01
 - (c) 0.4509
 - (d) 0.15
 - (e) 0.7564
36. A statistics course has three tutorials, TUT01, TUT02 and TUT03, each with 50 students. TUT01 has 20 girls and 30 boys, TUT02 has 30 girls and 20 boys and TUT03 has 40 girls and 10 boys. We are going to select a sample of three students from this course by selecting one student at random from each tutorial. What is the probability that our sample will have at least two girls?
- (a) 0.288
 - (b) 0.464
 - (c) 0.656
 - (d) 0.784
 - (e) 0.944

37. Historical evidence suggests that SAT scores are normally distributed with mean 1000 and standard deviation 180. What score (approximately) do you have to make to be in the top 1 % of all those who are taking this exam?
- (a) 586
 - (b) 1180
 - (c) 1234
 - (d) 1325
 - (e) 1415
38. The unemployment rate (i.e. the percent unemployed in the labour force) in a certain city is 8%. A random sample of 150 people from the labour force in this city is drawn. Find the approximate probability that this sample contains fifteen or more unemployed people. Which of the following numbers is closest to this probability?
- (a) 0.1
 - (b) 0.2
 - (c) 0.3
 - (d) 0.4
 - (e) 0.5
39. Consider the random experiment of rolling a six sided die for which the sample space is $S = \{1, 2, 3, 4, 5, 6\}$. We know that each side of the die has probability $1/6$ of coming up. Which of the following pairs of events are independent?
- (a) $A = \{1, 2\}$, $B = \{1, 2, 3\}$
 - (b) $A = \{1, 2\}$, $B = \{2, 3\}$
 - (c) $A = \{1, 3, 5\}$, $B = \{2, 4, 6\}$
 - (d) $A = \{1, 2, 3\}$, $B = \{3, 4\}$
 - (e) None of the above
40. In a certain game of chance, your chances of winning are 0.2 and it costs \$1 to play the game. If you win, you receive \$4 (for a net gain of \$3). If you lose, you receive nothing (for a net loss of \$1). You are going to play the game nine times. Assume that the outcomes are independent. Let T be the total net gain from this round of nine games. Find the standard deviation of T . Choose the closest.
- (a) \$ 1
 - (b) \$ 3
 - (c) \$ 5
 - (d) \$ 7
 - (e) \$ 9

41. A random sample of size 4 is taken from a large (infinite) population with mean 100 and standard deviation 20 respectively. The mean and the standard deviation of the distribution of the sample mean are:
- 100 and 20 respectively
 - 50 and 10 respectively
 - 50 and 20 respectively
 - 100 and 10 respectively
 - none of the above answers is correct
42. In a study, a group of researchers measured the blood pressure levels of 474 patients, recording the blood pressure levels as high blood pressure (HBP), low blood pressure (LBP) or normal blood pressure (NBP). The patients were also classified into three age categories: under 30, 30-50, and over 50. The results of this study are summarized in the table below:

Blood pressure level	Under 30	30-50	Over 50	Total
HBP	23	51	73	147
LBP	27	37	31	95
NBP	48	91	93	232
Total	98	179	197	474

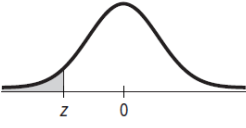
Which pie chart below represents the marginal distribution of blood pressure level?



- pie chart A
- pie chart B
- pie chart C
- pie chart D
- pie chart E

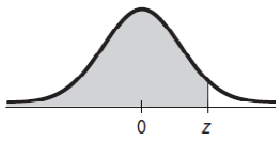
43. There are ten children in a room. The mean and standard deviation of their ages are 5 years and 1.25 years respectively. If another 5-year-old child enters this room, what will happen to the mean and standard deviation of the ages of the children in the room?
- (a) The mean will stay the same but the standard deviation will increase.
 - (b) The mean and standard deviation will both increase.
 - (c) The mean and standard deviation will both decrease.
 - (d) The mean will stay the same but the standard deviation will decrease.
 - (e) The mean and standard deviation will both stay the same.
44. In a large national population of post-secondary students, 60 percent attend 4-year institutions while the remainder attends 2-year institutions. Males make up 42 percent of post-secondary students. In 4-year institutions, 45 percent of the students are male. We will select one student at random. What is the probability the selected student is a male attending a 4-year institution?
- (a) 0.75
 - (b) 0.25
 - (c) 0.19
 - (d) 0.27
 - (e) 0.64

END OF EXAM

<p>Table Z Areas under the standard Normal curve</p> 		Second decimal place in z										
		0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.00	z
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0000 [†]	-3.9
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	-3.8
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	-3.7
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	-3.6
0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	-3.5
0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	-3.4
0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005	0.0005	-3.3
0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007	0.0007	-3.2
0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0009	0.0010	-3.1
0.0010	0.0010	0.0011	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013	0.0013	-3.0
0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0016	0.0017	0.0018	0.0018	0.0018	0.0019	-2.9
0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0023	0.0024	0.0025	0.0025	0.0026	-2.8
0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0031	0.0032	0.0033	0.0034	0.0034	0.0035	-2.7
0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0041	0.0043	0.0044	0.0045	0.0045	0.0047	-2.6
0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0055	0.0057	0.0059	0.0060	0.0060	0.0062	-2.5
0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0073	0.0075	0.0078	0.0080	0.0080	0.0082	-2.4
0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0096	0.0099	0.0102	0.0104	0.0104	0.0107	-2.3
0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0125	0.0129	0.0132	0.0136	0.0136	0.0139	-2.2
0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0162	0.0166	0.0170	0.0174	0.0174	0.0179	-2.1
0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0207	0.0212	0.0217	0.0222	0.0222	0.0228	-2.0
0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0262	0.0268	0.0274	0.0281	0.0281	0.0287	-1.9
0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0329	0.0336	0.0344	0.0351	0.0351	0.0359	-1.8
0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0409	0.0418	0.0427	0.0436	0.0436	0.0446	-1.7
0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0505	0.0516	0.0526	0.0537	0.0537	0.0548	-1.6
0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0618	0.0630	0.0643	0.0655	0.0655	0.0668	-1.5
0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0749	0.0764	0.0778	0.0793	0.0793	0.0808	-1.4
0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0901	0.0918	0.0934	0.0951	0.0951	0.0968	-1.3
0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1075	0.1093	0.1112	0.1131	0.1131	0.1151	-1.2
0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1271	0.1292	0.1314	0.1335	0.1335	0.1357	-1.1
0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1492	0.1515	0.1539	0.1562	0.1562	0.1587	-1.0
0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1736	0.1762	0.1788	0.1814	0.1814	0.1841	-0.9
0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2005	0.2033	0.2061	0.2090	0.2090	0.2119	-0.8
0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2296	0.2327	0.2358	0.2389	0.2389	0.2420	-0.7
0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2611	0.2643	0.2676	0.2709	0.2709	0.2743	-0.6
0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2946	0.2981	0.3015	0.3050	0.3050	0.3085	-0.5
0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3300	0.3336	0.3372	0.3409	0.3409	0.3446	-0.4
0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3669	0.3707	0.3745	0.3783	0.3783	0.3821	-0.3
0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4052	0.4090	0.4129	0.4168	0.4168	0.4207	-0.2
0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4443	0.4483	0.4522	0.4562	0.4562	0.4602	-0.1
0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4840	0.4880	0.4920	0.4960	0.4960	0.5000	-0.0

[†]For $z \leq -3.90$, the areas are 0.0000 to four decimal places.

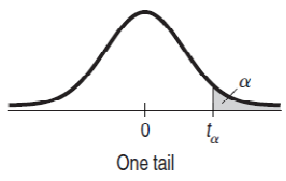
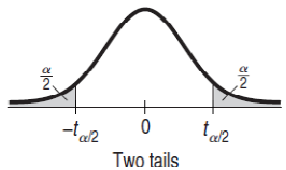
Table Z (cont.)
Areas under the
standard Normal curve



z	Second decimal place in z									
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000 [†]									

[†]For $z \geq 3.90$, the areas are 1.0000 to four decimal places.

		0.20	0.10	0.05	0.02	0.01	
Two-tail probability							
One-tail probability		0.10	0.05	0.025	0.01	0.005	
Table T	df						df
Values of t_α	1	3.078	6.314	12.706	31.821	63.657	1
	2	1.886	2.920	4.303	6.965	9.925	2
	3	1.638	2.353	3.182	4.541	5.841	3
	4	1.533	2.132	2.776	3.747	4.604	4
	5	1.476	2.015	2.571	3.365	4.032	5
	6	1.440	1.943	2.447	3.143	3.707	6
	7	1.415	1.895	2.365	2.998	3.499	7
	8	1.397	1.860	2.306	2.896	3.355	8
	9	1.383	1.833	2.262	2.821	3.250	9
	10	1.372	1.812	2.228	2.764	3.169	10
	11	1.363	1.796	2.201	2.718	3.106	11
	12	1.356	1.782	2.179	2.681	3.055	12
	13	1.350	1.771	2.160	2.650	3.012	13
	14	1.345	1.761	2.145	2.624	2.977	14
	15	1.341	1.753	2.131	2.602	2.947	15
	16	1.337	1.746	2.120	2.583	2.921	16
	17	1.333	1.740	2.110	2.567	2.898	17
	18	1.330	1.734	2.101	2.552	2.878	18
	19	1.328	1.729	2.093	2.539	2.861	19
	20	1.325	1.725	2.086	2.528	2.845	20
	21	1.323	1.721	2.080	2.518	2.831	21
	22	1.321	1.717	2.074	2.508	2.819	22
	23	1.319	1.714	2.069	2.500	2.807	23
	24	1.318	1.711	2.064	2.492	2.797	24
	25	1.316	1.708	2.060	2.485	2.787	25
	26	1.315	1.706	2.056	2.479	2.779	26
	27	1.314	1.703	2.052	2.473	2.771	27
	28	1.313	1.701	2.048	2.467	2.763	28
	29	1.311	1.699	2.045	2.462	2.756	29
	30	1.310	1.697	2.042	2.457	2.750	30
	32	1.309	1.694	2.037	2.449	2.738	32
	35	1.306	1.690	2.030	2.438	2.725	35
	40	1.303	1.684	2.021	2.423	2.704	40
	45	1.301	1.679	2.014	2.412	2.690	45
	50	1.299	1.676	2.009	2.403	2.678	50
	60	1.296	1.671	2.000	2.390	2.660	60
	75	1.293	1.665	1.992	2.377	2.643	75
	100	1.290	1.660	1.984	2.364	2.626	100
	120	1.289	1.658	1.980	2.358	2.617	120
	140	1.288	1.656	1.977	2.353	2.611	140
	180	1.286	1.653	1.973	2.347	2.603	180
	250	1.285	1.651	1.969	2.341	2.596	250
	400	1.284	1.649	1.966	2.336	2.588	400
	1000	1.282	1.646	1.962	2.330	2.581	1000
	∞	1.282	1.645	1.960	2.326	2.576	∞
Confidence levels		80%	90%	95%	98%	99%	



Binomial Distribution Table

n	k	0.1	0.2	0.3	0.4	0.5	0.6
14	0	0.2288	0.0440	0.0068	8e-04	1e-04	0.0000
14	1	0.3559	0.1539	0.0407	0.0073	9e-04	1e-04
14	2	0.2570	0.2501	0.1134	0.0317	0.0056	5e-04
14	3	0.1142	0.2501	0.1943	0.0845	0.0222	0.0033
14	4	0.0349	0.1720	0.2290	0.1549	0.0611	0.0136
14	5	0.0078	0.0860	0.1963	0.2066	0.1222	0.0408
14	6	0.0013	0.0322	0.1262	0.2066	0.1833	0.0918
14	7	2e-04	0.0092	0.0618	0.1574	0.2095	0.1574
14	8	0.0000	0.0020	0.0232	0.0918	0.1833	0.2066
14	9	0.0000	3e-04	0.0066	0.0408	0.1222	0.2066
14	10	0.0000	0.0000	0.0014	0.0136	0.0611	0.1549
14	11	0.0000	0.0000	2e-04	0.0033	0.0222	0.0845
14	12	0.0000	0.0000	0.0000	5e-04	0.0056	0.0317
14	13	0.0000	0.0000	0.0000	1e-04	9e-04	0.0073
14	14	0.0000	0.0000	0.0000	0.0000	1e-04	8e-04

n	k	0.1	0.2	0.3	0.4	0.5	0.6
15	0	0.2059	0.0352	0.0047	5e-04	0.0000	0.0000
15	1	0.3432	0.1319	0.0305	0.0047	5e-04	0.0000
15	2	0.2669	0.2309	0.0916	0.0219	0.0032	3e-04
15	3	0.1285	0.2501	0.1700	0.0634	0.0139	0.0016
15	4	0.0428	0.1876	0.2186	0.1268	0.0417	0.0074
15	5	0.0105	0.1032	0.2061	0.1859	0.0916	0.0245
15	6	0.0019	0.0430	0.1472	0.2066	0.1527	0.0612
15	7	3e-04	0.0138	0.0811	0.1771	0.1964	0.1181
15	8	0.0000	0.0035	0.0348	0.1181	0.1964	0.1771
15	9	0.0000	7e-04	0.0116	0.0612	0.1527	0.2066
15	10	0.0000	1e-04	0.0030	0.0245	0.0916	0.1859
15	11	0.0000	0.0000	6e-04	0.0074	0.0417	0.1268
15	12	0.0000	0.0000	1e-04	0.0016	0.0139	0.0634
15	13	0.0000	0.0000	0.0000	3e-04	0.0032	0.0219
15	14	0.0000	0.0000	0.0000	0.0000	5e-04	0.0047
15	15	0.0000	0.0000	0.0000	0.0000	0.0000	5e-04

UNIVERSITY OF TORONTO SCARBOROUGH
Department of Computer and Mathematical Sciences
DECEMBER 2013 EXAMINATIONS

STAB22H3 Statistics I

Duration: 3 hours

Last Name: _____ First Name: _____

Student number: _____

Aids allowed:

- Two handwritten letter-sized sheets (both sides) of notes prepared by you
- Non-programmable, non-communicating calculator

Standard Normal, t and the binomial distribution tables are attached at the end.

This test is based on multiple-choice questions. There are 44 questions. All questions carry equal weight. On the Scantron answer sheet, ensure that you enter your last name, first name (as much of it as fits), and student number (in “Identification”).

Mark in each case the best answer out of the alternatives given (**which means the numerically closest answer if the answer is a number and the answer you obtained is not given.**)

Also before you begin, complete the signature sheet, but sign it only when the invigilator collects it. The signature sheet shows that you were present at the exam.

Good luck!!

Answer Key for Exam A

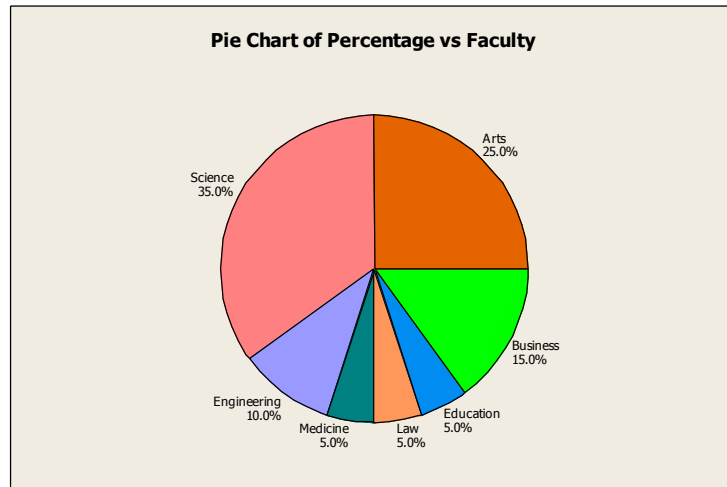
1. In a study of acupuncture for treating pain, 100 volunteers were recruited. Half were randomly assigned to receive acupuncture and the other half to receive a sham acupuncture treatment. The patients were followed for 6 months and the treating physician measured their degree of pain relief. The patients did not know which treatment they actually received, but the treating physicians were aware of who was getting acupuncture and who wasn't. Which of the following best describes this study?
 - (a) This is a retrospective observational study.
 - (b) This is a randomized block design.
 - (c) This is a randomized single-blind experiment.
 - (d) This is a randomized double-blind experiment.
 - (e) This is a prospective observational study.

2. Which of the following pairs of events (denoted by A and B) are disjoint and independent?
 - (a) Toss two coins, A = first toss is head, B = second toss is head
 - (b) Roll a die, A = observe an odd number, B = observe an even number
 - (c) Take two cards one by one from a deck of 52 cards, A = first card is red, B = both cards are red
 - (d) There are three kids in a family, A = two kids are girls, B = two kids are twins
 - (e) None of the above

3. In a group of 100 people, 40 own a cat, 25 own a dog, and 15 own a cat and a dog. Find the probability that a person chosen at random, owns a dog, given that he owns a cat.
 - (a) 0.375
 - (b) 0.6
 - (c) 0.25
 - (d) 0.4
 - (e) 0.15

4. Two stores (store A and store B) sell watermelons. The weights of watermelons at store A has a Normal distribution with mean 21 pounds and standard deviation 3 pounds. The weights of watermelons at store B has a Normal distribution with mean 18 pounds and standard deviation 2 pounds. The weights of watermelons at the two stores are independent. If you select a melon at random at each store, what is the probability that melon you selected at store B will be heavier? Choose the closest.
- (a) 0.1
 - (b) 0.2
 - (c) 0.3
 - (d) 0.4
 - (e) 0.5
5. The regression equation relating midterm score (x) and final exam score (y) in a course is $y = 20 + 0.8x$. Which of the following is a valid conclusion based on this information?
- (a) If Adam has scored 10 points higher than Bob on the midterm, then Adam will score 8 points higher than Bob on the final exam.
 - (b) Students who score 80 on the midterm are predicted to score 84 on the final.
 - (c) If a student scores 0 on the midterm, then that student will score 20 on the final.
 - (d) On average, the students' scores on the final exam will only be 80% of what they were on the midterm.
 - (e) This equation shows that the midterm and the final exam scores are not related.
6. A test of significance with $\alpha = 0.05$ results in a P-value of 0.1. What do you conclude?
- (a) Reject the null hypothesis.
 - (b) Reject the alternative hypothesis.
 - (c) Do not reject the alternative hypothesis.
 - (d) Do not reject the null hypothesis.
 - (e) Accept the null hypothesis.

7. The pie chart below shows the percentage of students in each faculty at a university.



If the number of students in the faculty of Arts is 3000, then how MANY students are there in the faculty of Science?

- (a) 3400
- (b) 3600
- (c) 3800
- (d) 4000
- (e) 4200

8. The two-way table below classifies the members of a fitness club by gender and food habits. Use this information to answer this question and the next question.

	Women	Men
Vegetarian	9	3
Non-vegetarian	8	10

If a member is selected at random from the club, what is the probability that this person is a vegetarian female?

- (a) $3/4$
- (b) $3/10$
- (c) $1/10$
- (d) $4/9$
- (e) $9/17$

9. Using the information in question 8 above, what is the probability that a randomly selected member is non-vegetarian?

- (a) 2/5
- (b) 1/3
- (c) 4/15
- (d) 3/5
- (e) 8/17

10. The regression equation and the summary statistics given below were obtained from a regression analysis of the relation between murder rate (y , the number of murders per 1,000,000 inhabitants per annum) and unemployment(x , the percentage of the labour force unemployed) using data for 20 cities.

Descriptive Statistics: x , y

Variable	N	Mean	StDev	Minimum	Q1	Q3	Maximum
x	20	deleted	1.207	4.900	6.050	8.125	9.300
y	20	20.57	9.88	5.30	12.88	26.63	40.70

Regression Analysis: y versus x

The regression equation is $y = - 28.5 + 7.08 x$

What is the correlation between x and y ?

- (a) 0.035
- (b) 0.0708
- (c) 0.122
- (d) 0.865
- (e) 0.988

11. Using the information in question 10 above, calculate the average unemployment rate (i.e. the value of \bar{x}) of these 20 cities. Note: This value has been deleted in the summary statistics given in question 10 above but can be calculated from other information given.

Which of the following numbers is closest to this value?

- (a) 7
- (b) 14
- (c) 21
- (d) 28
- (e) 49

12. One of the cities in the data set used in question 10 above, had unemployment rate 6.0 (i.e. $x = 6.0$) and murder rate 14.5. What is the residual for this city?
- (a) 0.52
 - (b) -68.16
 - (c) -27.98
 - (d) 43
 - (e) -96.66
13. Which of the following procedures best describes a method to select a simple random sample?
- (a) Randomly select half of the sample from females and the remaining half from males.
 - (b) Select an individual from every fourth house on a street.
 - (c) Assign each individual in the population a unique number and use a computer or a random number table to randomly generate numbers for selection.
 - (d) Select every individual with a surname beginning with the letter S.
 - (e) Select every 20th individual from a list of patients registered with a GP.
14. X and Y are two independent random variables and $Var(X) = 0.61$ and $Var(Y) = 0.21$. Find $SD(X - Y)$ (i.e. σ_{X-Y}).
- (a) 0.6451
 - (b) 0.6324
 - (c) 0.9055
 - (d) 0.5727
 - (e) 0.3227
15. In a large mall a survey was taken. It was found that in a random sample of 45 women over the age of 25, 15 had children. Find the 90% confidence interval for the population proportion of women over the age of 25 in the mall who have children. (Choose the closest answer.)
- (a) (0.0208, 0.0653)
 - (b) (0.4337, 0.6774)
 - (c) (0.3333, 0.5555)
 - (d) (0.1374, 0.1626)
 - (e) (0.2177, 0.4489)

16. In a regression study, which of the following best describes the square of the correlation r^2 between the two variables?
- (a) The amount of error that is incurred from using a regression model to make predictions
 - (b) The amount of the variability explained by the standard deviations of the response and explanatory variables
 - (c) The proportion of the total variation in the response variable explained by the straight line relationship with the explanatory variable
 - (d) The sum of squares of the residuals
 - (e) None of the above
17. Two independent random variables have the following distributions:

x	-1	0	1
p	0.2	0.3	0.5

y	1	2
p	0.3	0.7

Find the mean of $Y - 2X$ (i.e. find $E(Y - 2X)$).

- (a) 1.1
 - (b) -3.1
 - (c) 0.3
 - (d) 1.7
 - (e) 1.4
18. It is widely thought that there is a high incidence of disability among the homeless population. A random sample of 110 homeless people contained 84 who were disabled on one or more categories (such as psychiatric disability, medical disability etc.). Let p denote the proportion of all homeless people having one or more types of disability. Which of the following numbers is the closest to the value of the test Z-statistic for testing the null hypothesis $H_0 : p = 0.75$ against $H_a : p \neq 0.75$?
- (a) 0.0
 - (b) 0.3
 - (c) 0.6
 - (d) 0.9
 - (e) 1.2

19. The events A and B are such that $P(A) = 2P(B)$, $P(A \text{ and } B) = 0.08$ and A and B are independent. Find $P(A \text{ or } B)$.
- (a) 0.68
 - (b) 0.6
 - (c) 0.4
 - (d) 0.52
 - (e) 0
20. The heights of women in a certain population have a Normal distribution with mean 64.5 inches and standard deviation 2.5 inches. We select three women at random from this population. Assume that their heights are independent. Find the probability that the tallest of these three women will be taller than 67 inches. Which of the following numbers is closest to this probability?
- (a) 0.0
 - (b) 0.2
 - (c) 0.4
 - (d) 0.6
 - (e) 0.8
21. Which of the following values is closest to the interquartile range for the standard normal distribution?
- (a) 0
 - (b) 0.5
 - (c) 1.3
 - (d) 2.5
 - (e) 3

22. The distribution of bachelor's degrees conferred by a local college is listed below, by major.

Major	Frequency
English	2,073
Mathematics	2,164
Chemistry	318
Physics	856
Liberal Arts	1,358
Business	1,676
Engineering	868
Total	9,313

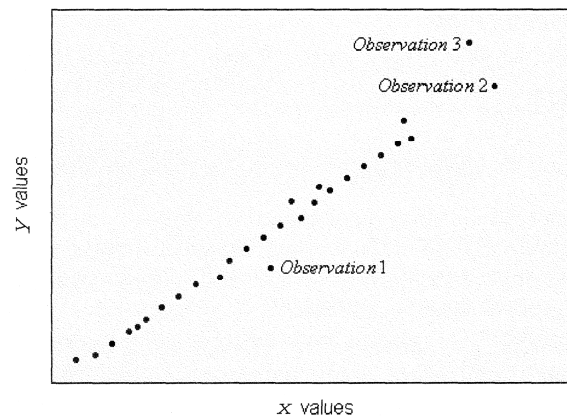
What is the probability that a randomly selected bachelor's degree is not in Mathematics?

- (a) 0.232
 (b) 0.768
(c) 0.303
(d) 0.682
(e) 0.889
23. You have sampled 25 students to find the mean SAT scores. A 95% confidence interval for the mean SAT score is 900 to 1100. Which of the following statements gives a valid interpretation of this interval?
- (a) 95% of the 25 students have a mean score between 900 and 1100.
(b) 95% of the population of all students have a score between 900 and 1100.
(c) If this procedure were repeated many times, 95% of the sample means would be between 900 and 1100.
 (d) If this procedure were repeated many times, 95% of the resulting confidence intervals would contain the true mean SAT score.
(e) If 100 samples were taken and a 95% confidence interval was computed, 5 of them would be in the interval from 900 to 1100.

24. A manufacturer of light bulbs claims that their light bulbs have a mean life of 1520 hours with an unknown standard deviation. A random sample of 30 such bulbs is selected for testing. You want to test the null hypothesis that the population mean life is less than the manufacturer claims. If the sample produces a mean value of 1505 hours and a sample standard deviation of 86, what can you conclude? You may assume that all the necessary assumptions for the related test are satisfied by the data.
- (a) We will reject the null hypothesis at $\alpha = 0.01$
 - (b) We will reject the null hypothesis at $\alpha = 0.05$
 - (c) We will reject the null hypothesis at $\alpha = 0.1$
 - (d) We will reject the null hypothesis at $\alpha = 0.001$
 - (e) We fail to reject the null hypothesis at $\alpha = 0.1$
25. The manufacturer of a bag of sweets claims that there is a 90% chance that the bag contains some toffees. If 20 bags are chosen, how many bags with toffees would you expect to have?
- (a) 20
 - (b) 2
 - (c) 9
 - (d) 18
 - (e) None
26. In brief, what does the Central Limit Theorem say?
- (a) The area under a Normal density curve is one.
 - (b) Measures of central tendency should always be computed with and without outliers.
 - (c) For sufficiently large sample size, the sampling distribution of \bar{X} is approximately Normal.
 - (d) Confidence intervals have zero margin of error for large sample sizes.
 - (e) In the long run, the average outcome gets close to the distribution mean.
27. A consumer product agency tests mileage (miles per gallon) for a sample of automobiles using each of four different types of gasoline. Which of the following statements regarding this study is true?
- (a) There are four explanatory variables and one response variable in this study.
 - (b) There is only one explanatory variable with four levels and one response variable in this study.
 - (c) Miles per gallon is the only explanatory variable in this study.
 - (d) This study has a single response variable with four levels.
 - (e) None of the above statements is true.

28. The heights (y) of 50 men and their shoes sizes (x) were obtained. The variable height is measured in centimetres (cm) and the shoe sizes of these men ranged from 8 to 13. From these 50 pairs of observations, the least squares regression line predicting height from shoe size was computed to be $\hat{y} = 130.455 + 4.7498x$. What height would you predict for a man with a shoe size of 13?
- (a) 130.46 cm
 - (b) 192.20 cm
 - (c) 182.70 cm
 - (d) This regression line cannot be used to predict the height of a man with a shoe size of 13
 - (e) None of the above
29. The weekly amount of money spent on maintenance and repairs by a company has a Normal distribution with mean \$400 and standard deviation \$ 20. If \$ 450 is budgeted for next week, what is the probability that the actual cost will exceed the budget? Choose the value closest to this probability.
- (a) 0.002
 - (b) 0.006
 - (c) 0.025
 - (d) 0.01
 - (e) 0.5
30. The exam scores in a class has a Normal distribution. 2.5 percent of the class failed (i.e. scored below 50) and 16 percent of the class got A's. A is the highest grade in this exam and a score of 86 or above qualifies for an A grade. Based on 68-95-99.7 % rule, what is the mean of this Normal distribution?
- (a) 56
 - (b) 62
 - (c) 68
 - (d) 74
 - (e) 78

31. Government data show that 10% of males under age 25 are unemployed. A random sample is taken of 400 males who are in the labor force and under age 25. Find the probability that the sample unemployment rate is 0.12 or more. (Hint: Use Normal approximation)
- (a) 0.1023
 (b) 0.0918
 (c) 0.9082
 (d) 0.5470
 (e) 0.4530
32. The scatterplot of two variables x and y is shown below. This scatterplot shows three observations that appear to be somewhat unusual.



Read the following statements labeled I, II, III and IV, carefully.

- I Observations 1 and 2 are not influential.
 II Observations 2 and 3 have high leverage.
 III Observations 2 and 3 have large residuals.
 IV Out of the three observations, only observation 3 is an outlier.

Which of the above statements are true?

- (a) Only I is true.
 (b) Only I and II are true.
 (c) only I, II, and III are true.
 (d) only IV is true.
 (e) All are true.

33. A number calculated with complete population data and quantifies a characteristic of the population is called ...
- (a) Datum
 - (b) Statistics
 - (c) Population
 - (d) Stratum
 - (e) Parameter
34. A missile protection system consists of 3 radar sets operating independently, each with probability 0.9 of detecting a missile entering the zone that is covered by the system. If a missile enters the zone, what is the probability that at least one radar set will detect the missile?
- (a) 0.9
 - (b) 0.729
 - (c) 0.271
 - (d) 0.99
 - (e) 0.999
35. In clinical trials a certain drug has a 10% success rate of curing a known disease. If 15 people are known to have the disease. What is the probability of at least 2 being cured?
- (a) 0.5491
 - (b) 0.01
 - (c) 0.4509
 - (d) 0.15
 - (e) 0.7564
36. A statistics course has three tutorials, TUT01, TUT02 and TUT03, each with 50 students. TUT01 has 20 girls and 30 boys, TUT02 has 30 girls and 20 boys and TUT03 has 40 girls and 10 boys. We are going to select a sample of three students from this course by selecting one student at random from each tutorial. What is the probability that our sample will have at least two girls?
- (a) 0.288
 - (b) 0.464
 - (c) 0.656
 - (d) 0.784
 - (e) 0.944

37. Historical evidence suggests that SAT scores are normally distributed with mean 1000 and standard deviation 180. What score (approximately) do you have to make to be in the top 1 % of all those who are taking this exam?
- (a) 586
 - (b) 1180
 - (c) 1234
 - (d) 1325
 - (e) 1415
38. The unemployment rate (i.e. the percent unemployed in the labour force) in a certain city is 8%. A random sample of 150 people from the labour force in this city is drawn. Find the approximate probability that this sample contains fifteen or more unemployed people. Which of the following numbers is closest to this probability?
- (a) 0.1
 - (b) 0.2
 - (c) 0.3
 - (d) 0.4
 - (e) 0.5
39. Consider the random experiment of rolling a six sided die for which the sample space is $S = \{1, 2, 3, 4, 5, 6\}$. We know that each side of the die has probability $1/6$ of coming up. Which of the following pairs of events are independent?
- (a) $A = \{1, 2\}$, $B = \{1, 2, 3\}$
 - (b) $A = \{1, 2\}$, $B = \{2, 3\}$
 - (c) $A = \{1, 3, 5\}$, $B = \{2, 4, 6\}$
 - (d) $A = \{1, 2, 3\}$, $B = \{3, 4\}$
 - (e) None of the above
40. In a certain game of chance, your chances of winning are 0.2 and it costs \$1 to play the game. If you win, you receive \$4 (for a net gain of \$3). If you lose, you receive nothing (for a net loss of \$1). You are going to play the game nine times. Assume that the outcomes are independent. Let T be the total net gain from this round of nine games. Find the standard deviation of T . Choose the closest.
- (a) \$ 1
 - (b) \$ 3
 - (c) \$ 5
 - (d) \$ 7
 - (e) \$ 9

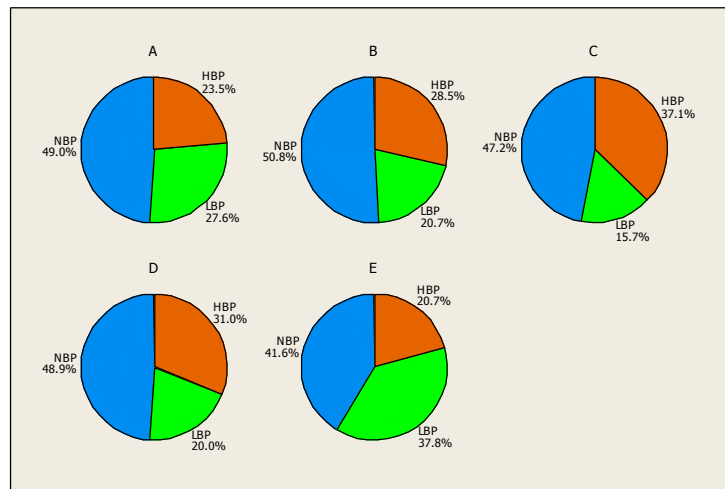
41. A random sample of size 4 is taken from a large (infinite) population with mean 100 and standard deviation 20 respectively. The mean and the standard deviation of the distribution of the sample mean are:

- (a) 100 and 20 respectively
- (b) 50 and 10 respectively
- (c) 50 and 20 respectively
- (d) 100 and 10 respectively
- (e) none of the above answers is correct

42. In a study, a group of researchers measured the blood pressure levels of 474 patients, recording the blood pressure levels as high blood pressure (HBP), low blood pressure (LBP) or normal blood pressure (NBP). The patients were also classified into three age categories: under 30, 30-50, and over 50. The results of this study are summarized in the table below:

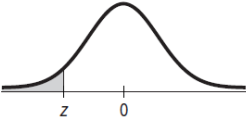
Blood pressure level	Under 30	30-50	Over 50	Total
HBP	23	51	73	147
LBP	27	37	31	95
NBP	48	91	93	232
Total	98	179	197	474

Which pie chart below represents the marginal distribution of blood pressure level?



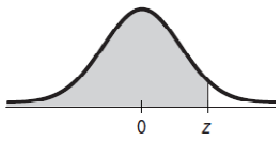
- (a) pie chart A
- (b) pie chart B
- (c) pie chart C
- (d) pie chart D
- (e) pie chart E

43. There are ten children in a room. The mean and standard deviation of their ages are 5 years and 1.25 years respectively. If another 5-year-old child enters this room, what will happen to the mean and standard deviation of the ages of the children in the room?
- (a) The mean will stay the same but the standard deviation will increase.
 - (b) The mean and standard deviation will both increase.
 - (c) The mean and standard deviation will both decrease.
 - (d) The mean will stay the same but the standard deviation will decrease.
 - (e) The mean and standard deviation will both stay the same.
44. In a large national population of post-secondary students, 60 percent attend 4-year institutions while the remainder attends 2-year institutions. Males make up 42 percent of post-secondary students. In 4-year institutions, 45 percent of the students are male. We will select one student at random. What is the probability the selected student is a male attending a 4-year institution?
- (a) 0.75
 - (b) 0.25
 - (c) 0.19
 - (d) 0.27
 - (e) 0.64

<p>Table Z Areas under the standard Normal curve</p> 		Second decimal place in z									z	
		0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01		0.00
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0000 [†]	-3.9
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	-3.8
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	-3.7
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	-3.6
0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	-3.5
0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	-3.4
0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005	0.0005	-3.3
0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007	0.0007	-3.2
0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0009	0.0010	-3.1
0.0010	0.0010	0.0011	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013	0.0013	-3.0
0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0016	0.0017	0.0018	0.0018	0.0018	0.0019	-2.9
0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0023	0.0024	0.0025	0.0025	0.0026	-2.8
0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0031	0.0032	0.0033	0.0034	0.0034	0.0035	-2.7
0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0041	0.0043	0.0044	0.0045	0.0045	0.0047	-2.6
0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0055	0.0057	0.0059	0.0060	0.0060	0.0062	-2.5
0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0073	0.0075	0.0078	0.0080	0.0080	0.0082	-2.4
0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0096	0.0099	0.0102	0.0104	0.0104	0.0107	-2.3
0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0125	0.0129	0.0132	0.0136	0.0136	0.0139	-2.2
0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0162	0.0166	0.0170	0.0174	0.0174	0.0179	-2.1
0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0207	0.0212	0.0217	0.0222	0.0222	0.0228	-2.0
0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0262	0.0268	0.0274	0.0281	0.0281	0.0287	-1.9
0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0329	0.0336	0.0344	0.0351	0.0351	0.0359	-1.8
0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0409	0.0418	0.0427	0.0436	0.0436	0.0446	-1.7
0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0505	0.0516	0.0526	0.0537	0.0537	0.0548	-1.6
0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0618	0.0630	0.0643	0.0655	0.0655	0.0668	-1.5
0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0749	0.0764	0.0778	0.0793	0.0793	0.0808	-1.4
0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0901	0.0918	0.0934	0.0951	0.0951	0.0968	-1.3
0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1075	0.1093	0.1112	0.1131	0.1131	0.1151	-1.2
0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1271	0.1292	0.1314	0.1335	0.1335	0.1357	-1.1
0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1492	0.1515	0.1539	0.1562	0.1562	0.1587	-1.0
0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1736	0.1762	0.1788	0.1814	0.1814	0.1841	-0.9
0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2005	0.2033	0.2061	0.2090	0.2090	0.2119	-0.8
0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2296	0.2327	0.2358	0.2389	0.2389	0.2420	-0.7
0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2611	0.2643	0.2676	0.2709	0.2709	0.2743	-0.6
0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2946	0.2981	0.3015	0.3050	0.3050	0.3085	-0.5
0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3300	0.3336	0.3372	0.3409	0.3409	0.3446	-0.4
0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3669	0.3707	0.3745	0.3783	0.3783	0.3821	-0.3
0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4052	0.4090	0.4129	0.4168	0.4168	0.4207	-0.2
0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4443	0.4483	0.4522	0.4562	0.4562	0.4602	-0.1
0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4840	0.4880	0.4920	0.4960	0.4960	0.5000	-0.0

[†]For $z \leq -3.90$, the areas are 0.0000 to four decimal places.

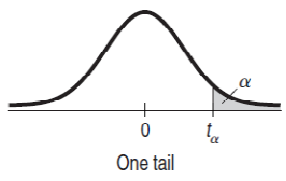
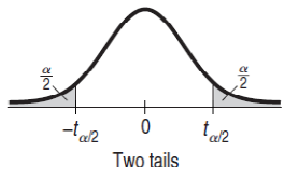
Table Z (cont.)
Areas under the
standard Normal curve



z	Second decimal place in z									
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000 [†]									

[†]For $z \geq 3.90$, the areas are 1.0000 to four decimal places.

Two-tail probability	0.20	0.10	0.05	0.02	0.01	
One-tail probability	0.10	0.05	0.025	0.01	0.005	
Table T						
df						df
Values of t_α						
1	3.078	6.314	12.706	31.821	63.657	1
2	1.886	2.920	4.303	6.965	9.925	2
3	1.638	2.353	3.182	4.541	5.841	3
4	1.533	2.132	2.776	3.747	4.604	4
5	1.476	2.015	2.571	3.365	4.032	5
6	1.440	1.943	2.447	3.143	3.707	6
7	1.415	1.895	2.365	2.998	3.499	7
8	1.397	1.860	2.306	2.896	3.355	8
9	1.383	1.833	2.262	2.821	3.250	9
10	1.372	1.812	2.228	2.764	3.169	10
11	1.363	1.796	2.201	2.718	3.106	11
12	1.356	1.782	2.179	2.681	3.055	12
13	1.350	1.771	2.160	2.650	3.012	13
14	1.345	1.761	2.145	2.624	2.977	14
15	1.341	1.753	2.131	2.602	2.947	15
16	1.337	1.746	2.120	2.583	2.921	16
17	1.333	1.740	2.110	2.567	2.898	17
18	1.330	1.734	2.101	2.552	2.878	18
19	1.328	1.729	2.093	2.539	2.861	19
20	1.325	1.725	2.086	2.528	2.845	20
21	1.323	1.721	2.080	2.518	2.831	21
22	1.321	1.717	2.074	2.508	2.819	22
23	1.319	1.714	2.069	2.500	2.807	23
24	1.318	1.711	2.064	2.492	2.797	24
25	1.316	1.708	2.060	2.485	2.787	25
26	1.315	1.706	2.056	2.479	2.779	26
27	1.314	1.703	2.052	2.473	2.771	27
28	1.313	1.701	2.048	2.467	2.763	28
29	1.311	1.699	2.045	2.462	2.756	29
30	1.310	1.697	2.042	2.457	2.750	30
32	1.309	1.694	2.037	2.449	2.738	32
35	1.306	1.690	2.030	2.438	2.725	35
40	1.303	1.684	2.021	2.423	2.704	40
45	1.301	1.679	2.014	2.412	2.690	45
50	1.299	1.676	2.009	2.403	2.678	50
60	1.296	1.671	2.000	2.390	2.660	60
75	1.293	1.665	1.992	2.377	2.643	75
100	1.290	1.660	1.984	2.364	2.626	100
120	1.289	1.658	1.980	2.358	2.617	120
140	1.288	1.656	1.977	2.353	2.611	140
180	1.286	1.653	1.973	2.347	2.603	180
250	1.285	1.651	1.969	2.341	2.596	250
400	1.284	1.649	1.966	2.336	2.588	400
1000	1.282	1.646	1.962	2.330	2.581	1000
∞	1.282	1.645	1.960	2.326	2.576	∞
Confidence levels	80%	90%	95%	98%	99%	



Binomial Distribution Table

n	k	0.1	0.2	0.3	0.4	0.5	0.6
14	0	0.2288	0.0440	0.0068	8e-04	1e-04	0.0000
14	1	0.3559	0.1539	0.0407	0.0073	9e-04	1e-04
14	2	0.2570	0.2501	0.1134	0.0317	0.0056	5e-04
14	3	0.1142	0.2501	0.1943	0.0845	0.0222	0.0033
14	4	0.0349	0.1720	0.2290	0.1549	0.0611	0.0136
14	5	0.0078	0.0860	0.1963	0.2066	0.1222	0.0408
14	6	0.0013	0.0322	0.1262	0.2066	0.1833	0.0918
14	7	2e-04	0.0092	0.0618	0.1574	0.2095	0.1574
14	8	0.0000	0.0020	0.0232	0.0918	0.1833	0.2066
14	9	0.0000	3e-04	0.0066	0.0408	0.1222	0.2066
14	10	0.0000	0.0000	0.0014	0.0136	0.0611	0.1549
14	11	0.0000	0.0000	2e-04	0.0033	0.0222	0.0845
14	12	0.0000	0.0000	0.0000	5e-04	0.0056	0.0317
14	13	0.0000	0.0000	0.0000	1e-04	9e-04	0.0073
14	14	0.0000	0.0000	0.0000	0.0000	1e-04	8e-04

n	k	0.1	0.2	0.3	0.4	0.5	0.6
15	0	0.2059	0.0352	0.0047	5e-04	0.0000	0.0000
15	1	0.3432	0.1319	0.0305	0.0047	5e-04	0.0000
15	2	0.2669	0.2309	0.0916	0.0219	0.0032	3e-04
15	3	0.1285	0.2501	0.1700	0.0634	0.0139	0.0016
15	4	0.0428	0.1876	0.2186	0.1268	0.0417	0.0074
15	5	0.0105	0.1032	0.2061	0.1859	0.0916	0.0245
15	6	0.0019	0.0430	0.1472	0.2066	0.1527	0.0612
15	7	3e-04	0.0138	0.0811	0.1771	0.1964	0.1181
15	8	0.0000	0.0035	0.0348	0.1181	0.1964	0.1771
15	9	0.0000	7e-04	0.0116	0.0612	0.1527	0.2066
15	10	0.0000	1e-04	0.0030	0.0245	0.0916	0.1859
15	11	0.0000	0.0000	6e-04	0.0074	0.0417	0.1268
15	12	0.0000	0.0000	1e-04	0.0016	0.0139	0.0634
15	13	0.0000	0.0000	0.0000	3e-04	0.0032	0.0219
15	14	0.0000	0.0000	0.0000	0.0000	5e-04	0.0047
15	15	0.0000	0.0000	0.0000	0.0000	0.0000	5e-04

UNIVERSITY OF TORONTO SCARBOROUGH
Department of Computer and Mathematical Sciences
DECEMBER 2013 EXAMINATIONS

STAB22H3 Statistics I

Duration: 3 hours

Last Name: _____ First Name: _____

Student number: _____

Aids allowed:

- Two handwritten letter-sized sheets (both sides) of notes prepared by you
- Non-programmable, non-communicating calculator

Standard Normal, t and the binomial distribution tables are attached at the end.

This test is based on multiple-choice questions. There are 44 questions. All questions carry equal weight. On the Scantron answer sheet, ensure that you enter your last name, first name (as much of it as fits), and student number (in “Identification”).

Mark in each case the best answer out of the alternatives given (**which means the numerically closest answer if the answer is a number and the answer you obtained is not given.**)

Also before you begin, complete the signature sheet, but sign it only when the invigilator collects it. The signature sheet shows that you were present at the exam.

Good luck!!

ExamVersion: B

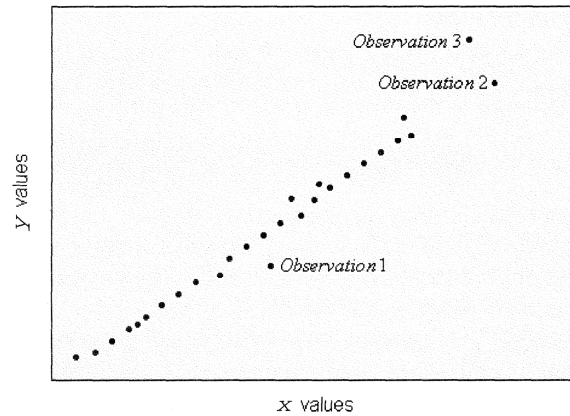
1. Consider the random experiment of rolling a six sided die for which the sample space is $S = \{1, 2, 3, 4, 5, 6\}$. We know that each side of the die has probability $1/6$ of coming up. Which of the following pairs of events are independent?
 - (a) $A = \{1, 2\}$, $B = \{1, 2, 3\}$
 - (b) $A = \{1, 2\}$, $B = \{2, 3\}$
 - (c) $A = \{1, 3, 5\}$, $B = \{2, 4, 6\}$
 - (d) $A = \{1, 2, 3\}$, $B = \{3, 4\}$
 - (e) None of the above

2. The heights (y) of 50 men and their shoes sizes (x) were obtained. The variable height is measured in centimetres (cm) and the shoe sizes of these men ranged from 8 to 13. From these 50 pairs of observations, the least squares regression line predicting height from shoe size was computed to be $\hat{y} = 130.455 + 4.7498x$. What height would you predict for a man with a shoe size of 13?
 - (a) 130.46 cm
 - (b) 192.20 cm
 - (c) 182.70 cm
 - (d) This regression line cannot be used to predict the height of a man with a shoe size of 13
 - (e) None of the above

3. In a study of acupuncture for treating pain, 100 volunteers were recruited. Half were randomly assigned to receive acupuncture and the other half to receive a sham acupuncture treatment. The patients were followed for 6 months and the treating physician measured their degree of pain relief. The patients did not know which treatment they actually received, but the treating physicians were aware of who was getting acupuncture and who wasn't. Which of the following best describes this study?
 - (a) This is a retrospective observational study.
 - (b) This is a randomized block design.
 - (c) This is a randomized single-blind experiment.
 - (d) This is a randomized double-blind experiment.
 - (e) This is a prospective observational study.

4. Historical evidence suggests that SAT scores are normally distributed with mean 1000 and standard deviation 180. What score (approximately) do you have to make to be in the top 1 % of all those who are taking this exam?
- (a) 586
 - (b) 1180
 - (c) 1234
 - (d) 1325
 - (e) 1415
5. Which of the following values is closest to the interquartile range for the standard normal distribution?
- (a) 0
 - (b) 0.5
 - (c) 1.3
 - (d) 2.5
 - (e) 3
6. In a large national population of post-secondary students, 60 percent attend 4-year institutions while the remainder attends 2-year institutions. Males make up 42 percent of post-secondary students. In 4-year institutions, 45 percent of the students are male. We will select one student at random. What is the probability the selected student is a male attending a 4-year institution?
- (a) 0.75
 - (b) 0.25
 - (c) 0.19
 - (d) 0.27
 - (e) 0.64

7. The scatterplot of two variables x and y is shown below. This scatterplot shows three observations that appear to be somewhat unusual.



Read the following statements labeled I, II, III and IV, carefully.

- I Observations 1 and 2 are not influential.
- II Observations 2 and 3 have high leverage.
- III Observations 2 and 3 have large residuals.
- IV Out of the three observations, only observation 3 is an outlier.

Which of the above statements are true?

- (a) Only I is true.
 - (b) Only I and II are true.
 - (c) only I, II, and III are true.
 - (d) only IV is true.
 - (e) All are true.
8. Which of the following procedures best describes a method to select a simple random sample?
- (a) Randomly select half of the sample from females and the remaining half from males.
 - (b) Select an individual from every fourth house on a street.
 - (c) Assign each individual in the population a unique number and use a computer or a random number table to randomly generate numbers for selection.
 - (d) Select every individual with a surname beginning with the letter S.
 - (e) Select every 20th individual from a list of patients registered with a GP.

9. A manufacturer of light bulbs claims that their light bulbs have a mean life of 1520 hours with an unknown standard deviation. A random sample of 30 such bulbs is selected for testing. You want to test the null hypothesis that the population mean life is less than the manufacturer claims. If the sample produces a mean value of 1505 hours and a sample standard deviation of 86, what can you conclude? You may assume that all the necessary assumptions for the related test are satisfied by the data.
- (a) We will reject the null hypothesis at $\alpha = 0.01$
 - (b) We will reject the null hypothesis at $\alpha = 0.05$
 - (c) We will reject the null hypothesis at $\alpha = 0.1$
 - (d) We will reject the null hypothesis at $\alpha = 0.001$
 - (e) We fail to reject the null hypothesis at $\alpha = 0.1$
10. The weekly amount of money spent on maintenance and repairs by a company has a Normal distribution with mean \$400 and standard deviation \$ 20. If \$ 450 is budgeted for next week, what is the probability that the actual cost will exceed the budget? Choose the value closest to this probability.
- (a) 0.002
 - (b) 0.006
 - (c) 0.025
 - (d) 0.01
 - (e) 0.5
11. A random sample of size 4 is taken from a large (infinite) population with mean 100 and standard deviation 20 respectively. The mean and the standard deviation of the distribution of the sample mean are:
- (a) 100 and 20 respectively
 - (b) 50 and 10 respectively
 - (c) 50 and 20 respectively
 - (d) 100 and 10 respectively
 - (e) none of the above answers is correct

12. In a regression study, which of the following best describes the square of the correlation r^2 between the two variables?
- (a) The amount of error that is incurred from using a regression model to make predictions
 - (b) The amount of the variability explained by the standard deviations of the response and explanatory variables
 - (c) The proportion of the total variation in the response variable explained by the straight line relationship with the explanatory variable
 - (d) The sum of squares of the residuals
 - (e) None of the above
13. Which of the following pairs of events (denoted by A and B) are disjoint and independent?
- (a) Toss two coins, A = first toss is head, B = second toss is head
 - (b) Roll a die, A = observe an odd number, B = observe an even number
 - (c) Take two cards one by one from a deck of 52 cards, A = first card is red, B = both cards are red
 - (d) There are three kids in a family, A = two kids are girls, B = two kids are twins
 - (e) None of the above
14. A missile protection system consists of 3 radar sets operating independently, each with probability 0.9 of detecting a missile entering the zone that is covered by the system. If a missile enters the zone, what is the probability that at least one radar set will detect the missile?
- (a) 0.9
 - (b) 0.729
 - (c) 0.271
 - (d) 0.99
 - (e) 0.999

15. In a large mall a survey was taken. It was found that in a random sample of 45 women over the age of 25, 15 had children. Find the 90% confidence interval for the population proportion of women over the age of 25 in the mall who have children. (Choose the closest answer.)
- (a) (0.0208, 0.0653)
 - (b) (0.4337, 0.6774)
 - (c) (0.3333, 0.5555)
 - (d) (0.1374, 0.1626)
 - (e) (0.2177, 0.4489)
16. You have sampled 25 students to find the mean SAT scores. A 95% confidence interval for the mean SAT score is 900 to 1100. Which of the following statements gives a valid interpretation of this interval?
- (a) 95% of the 25 students have a mean score between 900 and 1100.
 - (b) 95% of the population of all students have a score between 900 and 1100.
 - (c) If this procedure were repeated many times, 95% of the sample means would be between 900 and 1100.
 - (d) If this procedure were repeated many times, 95% of the resulting confidence intervals would contain the true mean SAT score.
 - (e) If 100 samples were taken and a 95% confidence interval was computed, 5 of them would be in the interval from 900 to 1100.
17. A number calculated with complete population data and quantifies a characteristic of the population is called ...
- (a) Datum
 - (b) Statistics
 - (c) Population
 - (d) Stratum
 - (e) Parameter
18. In a group of 100 people, 40 own a cat, 25 own a dog, and 15 own a cat and a dog. Find the probability that a person chosen at random, owns a dog, given that he owns a cat.
- (a) 0.375
 - (b) 0.6
 - (c) 0.25
 - (d) 0.4
 - (e) 0.15

19. In clinical trials a certain drug has a 10% success rate of curing a known disease. If 15 people are known to have the disease. What is the probability of at least 2 being cured?
- (a) 0.5491
 - (b) 0.01
 - (c) 0.4509
 - (d) 0.15
 - (e) 0.7564
20. The exam scores in a class has a Normal distribution. 2.5 percent of the class failed (i.e. scored below 50) and 16 percent of the class got A's. A is the highest grade in this exam and a score of 86 or above qualifies for an A grade. Based on 68-95-99.7 % rule, what is the mean of this Normal distribution?
- (a) 56
 - (b) 62
 - (c) 68
 - (d) 74
 - (e) 78
21. In brief, what does the Central Limit Theorem say?
- (a) The area under a Normal density curve is one.
 - (b) Measures of central tendency should always be computed with and without outliers.
 - (c) For sufficiently large sample size, the sampling distribution of \bar{X} is approximately Normal.
 - (d) Confidence intervals have zero margin of error for large sample sizes.
 - (e) In the long run, the average outcome gets close to the distribution mean.
22. The events A and B are such that $P(A) = 2P(B)$, $P(A \text{ and } B) = 0.08$ and A and B are independent. Find $P(A \text{ or } B)$.
- (a) 0.68
 - (b) 0.6
 - (c) 0.4
 - (d) 0.52
 - (e) 0

23. In a certain game of chance, your chances of winning are 0.2 and it costs \$1 to play the game. If you win, you receive \$4 (for a net gain of \$3). If you lose, you receive nothing (for a net loss of \$1). You are going to play the game nine times. Assume that the outcomes are independent. Let T be the total net gain from this round of nine games. Find the standard deviation of T . Choose the closest.
- (a) \$ 1
 - (b) \$ 3
 - (c) \$ 5
 - (d) \$ 7
 - (e) \$ 9
24. The manufacturer of a bag of sweets claims that there is a 90% chance that the bag contains some toffees. If 20 bags are chosen, how many bags with toffees would you expect to have?
- (a) 20
 - (b) 2
 - (c) 9
 - (d) 18
 - (e) None
25. A statistics course has three tutorials, TUT01, TUT02 and TUT03, each with 50 students. TUT01 has 20 girls and 30 boys, TUT02 has 30 girls and 20 boys and TUT03 has 40 girls and 10 boys. We are going to select a sample of three students from this course by selecting one student at random from each tutorial. What is the probability that our sample will have at least two girls?
- (a) 0.288
 - (b) 0.464
 - (c) 0.656
 - (d) 0.784
 - (e) 0.944

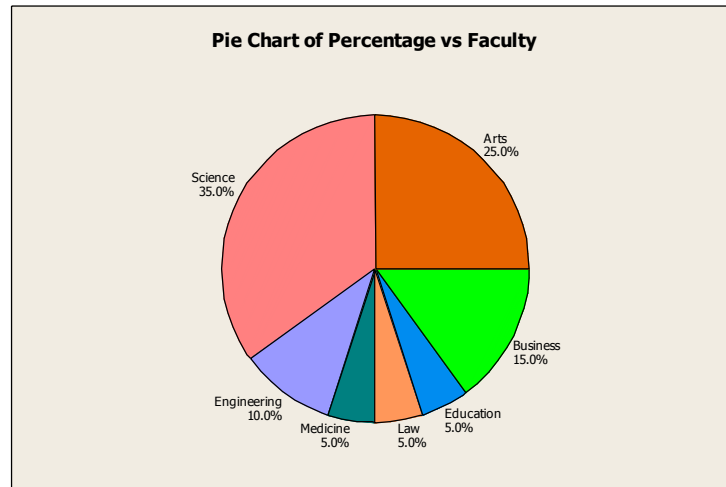
26. The distribution of bachelor's degrees conferred by a local college is listed below, by major.

Major	Frequency
English	2,073
Mathematics	2,164
Chemistry	318
Physics	856
Liberal Arts	1,358
Business	1,676
Engineering	868
Total	9,313

What is the probability that a randomly selected bachelor's degree is not in Mathematics?

- (a) 0.232
- (b) 0.768
- (c) 0.303
- (d) 0.682
- (e) 0.889

27. The pie chart below shows the percentage of students in each faculty at a university.



If the number of students in the faculty of Arts is 3000, then how MANY students are there in the faculty of Science?

- (a) 3400
 - (b) 3600
 - (c) 3800
 - (d) 4000
 - (e) 4200
28. Government data show that 10% of males under age 25 are unemployed. A random sample is taken of 400 males who are in the labor force and under age 25. Find the probability that the sample unemployment rate is 0.12 or more. (Hint: Use Normal approximation)
- (a) 0.1023
 - (b) 0.0918
 - (c) 0.9082
 - (d) 0.5470
 - (e) 0.4530

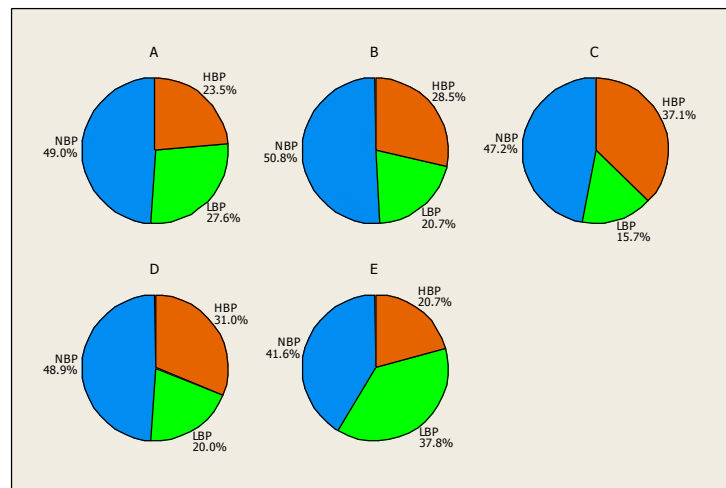
29. The unemployment rate (i.e. the percent unemployed in the labour force) in a certain city is 8%. A random sample of 150 people from the labour force in this city is drawn. Find the approximate probability that this sample contains fifteen or more unemployed people. Which of the following numbers is closest to this probability?

- (a) 0.1
- (b) 0.2
- (c) 0.3
- (d) 0.4
- (e) 0.5

30. In a study, a group of researchers measured the blood pressure levels of 474 patients, recording the blood pressure levels as high blood pressure (HBP), low blood pressure (LBP) or normal blood pressure (NBP). The patients were also classified into three age categories: under 30, 30-50, and over 50. The results of this study are summarized in the table below:

Blood pressure level	Under 30	30-50	Over 50	Total
HBP	23	51	73	147
LBP	27	37	31	95
NBP	48	91	93	232
Total	98	179	197	474

Which pie chart below represents the marginal distribution of blood pressure level?



- (a) pie chart A
 (b) pie chart B
 (c) pie chart C
 (d) pie chart D
 (e) pie chart E
31. A test of significance with $\alpha = 0.05$ results in a P-value of 0.1. What do you conclude?
- (a) Reject the null hypothesis.
 (b) Reject the alternative hypothesis.
 (c) Do not reject the alternative hypothesis.
 (d) Do not reject the null hypothesis.
 (e) Accept the null hypothesis.

32. There are ten children in a room. The mean and standard deviation of their ages are 5 years and 1.25 years respectively. If another 5-year-old child enters this room, what will happen to the mean and standard deviation of the ages of the children in the room?
- (a) The mean will stay the same but the standard deviation will increase.
 - (b) The mean and standard deviation will both increase.
 - (c) The mean and standard deviation will both decrease.
 - (d) The mean will stay the same but the standard deviation will decrease.
 - (e) The mean and standard deviation will both stay the same.
33. Two stores (store A and store B) sell watermelons. The weights of watermelons at store A has a Normal distribution with mean 21 pounds and standard deviation 3 pounds. The weights of watermelons at store B has a Normal distribution with mean 18 pounds and standard deviation 2 pounds. The weights of watermelons at the two stores are independent. If you select a melon at random at each store, what is the probability that melon you selected at store B will be heavier? Choose the closest.
- (a) 0.1
 - (b) 0.2
 - (c) 0.3
 - (d) 0.4
 - (e) 0.5
34. The regression equation relating midterm score (x) and final exam score (y) in a course is $y = 20 + 0.8x$. Which of the following is a valid conclusion based on this information?
- (a) If Adam has scored 10 points higher than Bob on the midterm, then Adam will score 8 points higher than Bob on the final exam.
 - (b) Students who score 80 on the midterm are predicted to score 84 on the final.
 - (c) If a student scores 0 on the midterm, then that student will score 20 on the final.
 - (d) On average, the students' scores on the final exam will only be 80% of what they were on the midterm.
 - (e) This equation shows that the midterm and the final exam scores are not related.

35. The heights of women in a certain population have a Normal distribution with mean 64.5 inches and standard deviation 2.5 inches. We select three women at random from this population. Assume that their heights are independent. Find the probability that the tallest of these three women will be taller than 67 inches. Which of the following numbers is closest to this probability?
- (a) 0.0
 - (b) 0.2
 - (c) 0.4
 - (d) 0.6
 - (e) 0.8
36. X and Y are two independent random variables and $Var(X) = 0.61$ and $Var(Y) = 0.21$. Find $SD(X - Y)$ (i.e. σ_{X-Y}).
- (a) 0.6451
 - (b) 0.6324
 - (c) 0.9055
 - (d) 0.5727
 - (e) 0.3227
37. A consumer product agency tests mileage (miles per gallon) for a sample of automobiles using each of four different types of gasoline. Which of the following statements regarding this study is true?
- (a) There are four explanatory variables and one response variable in this study.
 - (b) There is only one explanatory variable with four levels and one response variable in this study.
 - (c) Miles per gallon is the only explanatory variable in this study.
 - (d) This study has a single response variable with four levels.
 - (e) None of the above statements is true.

38. The two-way table below classifies the members of a fitness club by gender and food habits. Use this information to answer this question and the next question.

	Women	Men
Vegetarian	9	3
Non-vegetarian	8	10

- If a member is selected at random from the club, what is the probability that this person is a vegetarian female?
- (a) $3/4$
 - (b) $3/10$
 - (c) $1/10$
 - (d) $4/9$
 - (e) $9/17$
39. Using the information in question 38 above, what is the probability that a randomly selected member is non-vegetarian?
- (a) $2/5$
 - (b) $1/3$
 - (c) $4/15$
 - (d) $3/5$
 - (e) $8/17$
40. It is widely thought that there is a high incidence of disability among the homeless population. A random sample of 110 homeless people contained 84 who were disabled on one or more categories (such as psychiatric disability, medical disability etc.). Let p denote the proportion of all homeless people having one or more types of disability. Which of the following numbers is the closest to the value of the test Z-statistic for testing the null hypothesis $H_0 : p = 0.75$ against $H_a : p \neq 0.75$?
- (a) 0.0
 - (b) 0.3
 - (c) 0.6
 - (d) 0.9
 - (e) 1.2

41. Two independent random variables have the following distributions:

x	-1	0	1
p	0.2	0.3	0.5

y	1	2
p	0.3	0.7

Find the mean of $Y - 2X$ (i.e. find $E(Y - 2X)$).

- (a) 1.1
- (b) -3.1
- (c) 0.3
- (d) 1.7
- (e) 1.4

42. The regression equation and the summary statistics given below were obtained from a regression analysis of the relation between murder rate (y , the number of murders per 1,000,000 inhabitants per annum) and unemployment (x , the percentage of the labour force unemployed) using data for 20 cities.

Descriptive Statistics: x , y

Variable	N	Mean	StDev	Minimum	Q1	Q3	Maximum
x	20	deleted	1.207	4.900	6.050	8.125	9.300
y	20	20.57	9.88	5.30	12.88	26.63	40.70

Regression Analysis: y versus x

The regression equation is $y = - 28.5 + 7.08 x$

What is the correlation between x and y ?

- (a) 0.035
- (b) 0.0708
- (c) 0.122
- (d) 0.865
- (e) 0.988

43. Using the information in question 42 above, calculate the average unemployment rate (i.e. the value of \bar{x}) of these 20 cities. Note: This value has been deleted in the summary statistics given in question 42 above but can be calculated from other information given.

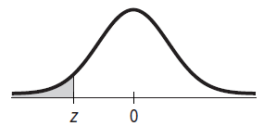
Which of the following numbers is closest to this value?

- (a) 7
- (b) 14
- (c) 21
- (d) 28
- (e) 49

44. One of the cities in the data set used in question 42 above, had unemployment rate 6.0 (i.e. $x = 6.0$) and murder rate 14.5. What is the residual for this city?

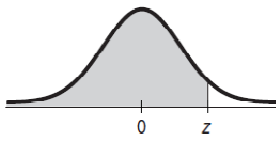
- (a) 0.52
- (b) -68.16
- (c) -27.98
- (d) 43
- (e) -96.66

END OF EXAM

<p>Table Z Areas under the standard Normal curve</p> 		Second decimal place in z									z	
		0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01		0.00
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0000 [†]	-3.9
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	-3.8
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	-3.7
0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	-3.6
0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	-3.5
0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	-3.4
0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005	0.0005	-3.3
0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007	0.0007	-3.2
0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0010	0.0010	-3.1
0.0010	0.0010	0.0011	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013	0.0013	-3.0
0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0016	0.0017	0.0018	0.0018	0.0019	0.0019	-2.9
0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0023	0.0024	0.0025	0.0026	0.0026	-2.8
0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0031	0.0032	0.0033	0.0034	0.0035	0.0035	-2.7
0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0041	0.0043	0.0044	0.0045	0.0047	0.0047	-2.6
0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0055	0.0057	0.0059	0.0060	0.0062	0.0062	-2.5
0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0073	0.0075	0.0078	0.0080	0.0082	0.0082	-2.4
0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0096	0.0099	0.0102	0.0104	0.0107	0.0107	-2.3
0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0125	0.0129	0.0132	0.0136	0.0139	0.0139	-2.2
0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0162	0.0166	0.0170	0.0174	0.0179	0.0179	-2.1
0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0207	0.0212	0.0217	0.0222	0.0228	0.0228	-2.0
0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0262	0.0268	0.0274	0.0281	0.0287	0.0287	-1.9
0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0329	0.0336	0.0344	0.0351	0.0359	0.0359	-1.8
0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0409	0.0418	0.0427	0.0436	0.0446	0.0446	-1.7
0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0505	0.0516	0.0526	0.0537	0.0548	0.0548	-1.6
0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0618	0.0630	0.0643	0.0655	0.0668	0.0668	-1.5
0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0749	0.0764	0.0778	0.0793	0.0808	0.0808	-1.4
0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0901	0.0918	0.0934	0.0951	0.0968	0.0968	-1.3
0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1075	0.1093	0.1112	0.1131	0.1151	0.1151	-1.2
0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1271	0.1292	0.1314	0.1335	0.1357	0.1357	-1.1
0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1492	0.1515	0.1539	0.1562	0.1587	0.1587	-1.0
0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1736	0.1762	0.1788	0.1814	0.1841	0.1841	-0.9
0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2005	0.2033	0.2061	0.2090	0.2119	0.2119	-0.8
0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2296	0.2327	0.2358	0.2389	0.2420	0.2420	-0.7
0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2611	0.2643	0.2676	0.2709	0.2743	0.2743	-0.6
0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2946	0.2981	0.3015	0.3050	0.3085	0.3085	-0.5
0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3300	0.3336	0.3372	0.3409	0.3446	0.3446	-0.4
0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3669	0.3707	0.3745	0.3783	0.3821	0.3821	-0.3
0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4052	0.4090	0.4129	0.4168	0.4207	0.4207	-0.2
0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4443	0.4483	0.4522	0.4562	0.4602	0.4602	-0.1
0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4840	0.4880	0.4920	0.4960	0.5000	0.5000	-0.0

[†]For $z \leq -3.90$, the areas are 0.0000 to four decimal places.

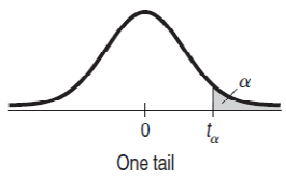
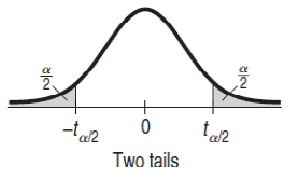
Table Z (cont.)
Areas under the
standard Normal curve



z	Second decimal place in z									
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000 [†]									

[†]For $z \geq 3.90$, the areas are 1.0000 to four decimal places.

Two-tail probability	0.20	0.10	0.05	0.02	0.01	
One-tail probability	0.10	0.05	0.025	0.01	0.005	
Table T						
df						df
Values of t_α						
1	3.078	6.314	12.706	31.821	63.657	1
2	1.886	2.920	4.303	6.965	9.925	2
3	1.638	2.353	3.182	4.541	5.841	3
4	1.533	2.132	2.776	3.747	4.604	4
5	1.476	2.015	2.571	3.365	4.032	5
6	1.440	1.943	2.447	3.143	3.707	6
7	1.415	1.895	2.365	2.998	3.499	7
8	1.397	1.860	2.306	2.896	3.355	8
9	1.383	1.833	2.262	2.821	3.250	9
10	1.372	1.812	2.228	2.764	3.169	10
11	1.363	1.796	2.201	2.718	3.106	11
12	1.356	1.782	2.179	2.681	3.055	12
13	1.350	1.771	2.160	2.650	3.012	13
14	1.345	1.761	2.145	2.624	2.977	14
15	1.341	1.753	2.131	2.602	2.947	15
16	1.337	1.746	2.120	2.583	2.921	16
17	1.333	1.740	2.110	2.567	2.898	17
18	1.330	1.734	2.101	2.552	2.878	18
19	1.328	1.729	2.093	2.539	2.861	19
20	1.325	1.725	2.086	2.528	2.845	20
21	1.323	1.721	2.080	2.518	2.831	21
22	1.321	1.717	2.074	2.508	2.819	22
23	1.319	1.714	2.069	2.500	2.807	23
24	1.318	1.711	2.064	2.492	2.797	24
25	1.316	1.708	2.060	2.485	2.787	25
26	1.315	1.706	2.056	2.479	2.779	26
27	1.314	1.703	2.052	2.473	2.771	27
28	1.313	1.701	2.048	2.467	2.763	28
29	1.311	1.699	2.045	2.462	2.756	29
30	1.310	1.697	2.042	2.457	2.750	30
32	1.309	1.694	2.037	2.449	2.738	32
35	1.306	1.690	2.030	2.438	2.725	35
40	1.303	1.684	2.021	2.423	2.704	40
45	1.301	1.679	2.014	2.412	2.690	45
50	1.299	1.676	2.009	2.403	2.678	50
60	1.296	1.671	2.000	2.390	2.660	60
75	1.293	1.665	1.992	2.377	2.643	75
100	1.290	1.660	1.984	2.364	2.626	100
120	1.289	1.658	1.980	2.358	2.617	120
140	1.288	1.656	1.977	2.353	2.611	140
180	1.286	1.653	1.973	2.347	2.603	180
250	1.285	1.651	1.969	2.341	2.596	250
400	1.284	1.649	1.966	2.336	2.588	400
1000	1.282	1.646	1.962	2.330	2.581	1000
∞	1.282	1.645	1.960	2.326	2.576	∞
Confidence levels	80%	90%	95%	98%	99%	



Binomial Distribution Table

n	k	0.1	0.2	0.3	0.4	0.5	0.6
14	0	0.2288	0.0440	0.0068	8e-04	1e-04	0.0000
14	1	0.3559	0.1539	0.0407	0.0073	9e-04	1e-04
14	2	0.2570	0.2501	0.1134	0.0317	0.0056	5e-04
14	3	0.1142	0.2501	0.1943	0.0845	0.0222	0.0033
14	4	0.0349	0.1720	0.2290	0.1549	0.0611	0.0136
14	5	0.0078	0.0860	0.1963	0.2066	0.1222	0.0408
14	6	0.0013	0.0322	0.1262	0.2066	0.1833	0.0918
14	7	2e-04	0.0092	0.0618	0.1574	0.2095	0.1574
14	8	0.0000	0.0020	0.0232	0.0918	0.1833	0.2066
14	9	0.0000	3e-04	0.0066	0.0408	0.1222	0.2066
14	10	0.0000	0.0000	0.0014	0.0136	0.0611	0.1549
14	11	0.0000	0.0000	2e-04	0.0033	0.0222	0.0845
14	12	0.0000	0.0000	0.0000	5e-04	0.0056	0.0317
14	13	0.0000	0.0000	0.0000	1e-04	9e-04	0.0073
14	14	0.0000	0.0000	0.0000	0.0000	1e-04	8e-04

n	k	0.1	0.2	0.3	0.4	0.5	0.6
15	0	0.2059	0.0352	0.0047	5e-04	0.0000	0.0000
15	1	0.3432	0.1319	0.0305	0.0047	5e-04	0.0000
15	2	0.2669	0.2309	0.0916	0.0219	0.0032	3e-04
15	3	0.1285	0.2501	0.1700	0.0634	0.0139	0.0016
15	4	0.0428	0.1876	0.2186	0.1268	0.0417	0.0074
15	5	0.0105	0.1032	0.2061	0.1859	0.0916	0.0245
15	6	0.0019	0.0430	0.1472	0.2066	0.1527	0.0612
15	7	3e-04	0.0138	0.0811	0.1771	0.1964	0.1181
15	8	0.0000	0.0035	0.0348	0.1181	0.1964	0.1771
15	9	0.0000	7e-04	0.0116	0.0612	0.1527	0.2066
15	10	0.0000	1e-04	0.0030	0.0245	0.0916	0.1859
15	11	0.0000	0.0000	6e-04	0.0074	0.0417	0.1268
15	12	0.0000	0.0000	1e-04	0.0016	0.0139	0.0634
15	13	0.0000	0.0000	0.0000	3e-04	0.0032	0.0219
15	14	0.0000	0.0000	0.0000	0.0000	5e-04	0.0047
15	15	0.0000	0.0000	0.0000	0.0000	0.0000	5e-04

UNIVERSITY OF TORONTO SCARBOROUGH
Department of Computer and Mathematical Sciences
DECEMBER 2013 EXAMINATIONS

STAB22H3 Statistics I

Duration: 3 hours

Last Name: _____ First Name: _____

Student number: _____

Aids allowed:

- Two handwritten letter-sized sheets (both sides) of notes prepared by you
- Non-programmable, non-communicating calculator

Standard Normal, t and the binomial distribution tables are attached at the end.

This test is based on multiple-choice questions. There are 44 questions. All questions carry equal weight. On the Scantron answer sheet, ensure that you enter your last name, first name (as much of it as fits), and student number (in “Identification”).

Mark in each case the best answer out of the alternatives given (**which means the numerically closest answer if the answer is a number and the answer you obtained is not given.**)

Also before you begin, complete the signature sheet, but sign it only when the invigilator collects it. The signature sheet shows that you were present at the exam.

Good luck!!

Answer Key for Exam **B**

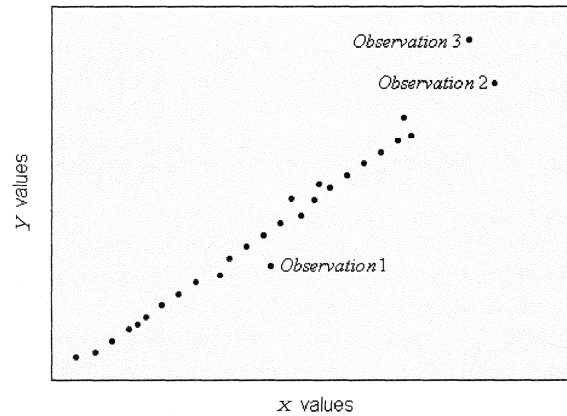
1. Consider the random experiment of rolling a six sided die for which the sample space is $S = \{1, 2, 3, 4, 5, 6\}$. We know that each side of the die has probability $1/6$ of coming up. Which of the following pairs of events are independent?
 - (a) $A = \{1, 2\}$, $B = \{1, 2, 3\}$
 - (b) $A = \{1, 2\}$, $B = \{2, 3\}$
 - (c) $A = \{1, 3, 5\}$, $B = \{2, 4, 6\}$
 - (d) $A = \{1, 2, 3\}$, $B = \{3, 4\}$
 - (e) None of the above

2. The heights (y) of 50 men and their shoes sizes (x) were obtained. The variable height is measured in centimetres (cm) and the shoe sizes of these men ranged from 8 to 13. From these 50 pairs of observations, the least squares regression line predicting height from shoe size was computed to be $\hat{y} = 130.455 + 4.7498x$. What height would you predict for a man with a shoe size of 13?
 - (a) 130.46 cm
 - (b) 192.20 cm
 - (c) 182.70 cm
 - (d) This regression line cannot be used to predict the height of a man with a shoe size of 13
 - (e) None of the above

3. In a study of acupuncture for treating pain, 100 volunteers were recruited. Half were randomly assigned to receive acupuncture and the other half to receive a sham acupuncture treatment. The patients were followed for 6 months and the treating physician measured their degree of pain relief. The patients did not know which treatment they actually received, but the treating physicians were aware of who was getting acupuncture and who wasn't. Which of the following best describes this study?
 - (a) This is a retrospective observational study.
 - (b) This is a randomized block design.
 - (c) This is a randomized single-blind experiment.
 - (d) This is a randomized double-blind experiment.
 - (e) This is a prospective observational study.

4. Historical evidence suggests that SAT scores are normally distributed with mean 1000 and standard deviation 180. What score (approximately) do you have to make to be in the top 1 % of all those who are taking this exam?
- (a) 586
 - (b) 1180
 - (c) 1234
 - (d) 1325
 - (e) 1415
5. Which of the following values is closest to the interquartile range for the standard normal distribution?
- (a) 0
 - (b) 0.5
 - (c) 1.3
 - (d) 2.5
 - (e) 3
6. In a large national population of post-secondary students, 60 percent attend 4-year institutions while the remainder attends 2-year institutions. Males make up 42 percent of post-secondary students. In 4-year institutions, 45 percent of the students are male. We will select one student at random. What is the probability the selected student is a male attending a 4-year institution?
- (a) 0.75
 - (b) 0.25
 - (c) 0.19
 - (d) 0.27
 - (e) 0.64

7. The scatterplot of two variables x and y is shown below. This scatterplot shows three observations that appear to be somewhat unusual.



Read the following statements labeled I, II, III and IV, carefully.

- I Observations 1 and 2 are not influential.
- II Observations 2 and 3 have high leverage.
- III Observations 2 and 3 have large residuals.
- IV Out of the three observations, only observation 3 is an outlier.

Which of the above statements are true?

- (a) Only I is true.
 - (b) Only I and II are true.
 - (c) only I, II, and III are true.
 - (d) only IV is true.
 - (e) All are true.
8. Which of the following procedures best describes a method to select a simple random sample?
- (a) Randomly select half of the sample from females and the remaining half from males.
 - (b) Select an individual from every fourth house on a street.
 - (c) Assign each individual in the population a unique number and use a computer or a random number table to randomly generate numbers for selection.
 - (d) Select every individual with a surname beginning with the letter S.
 - (e) Select every 20th individual from a list of patients registered with a GP.

9. A manufacturer of light bulbs claims that their light bulbs have a mean life of 1520 hours with an unknown standard deviation. A random sample of 30 such bulbs is selected for testing. You want to test the null hypothesis that the population mean life is less than the manufacturer claims. If the sample produces a mean value of 1505 hours and a sample standard deviation of 86, what can you conclude? You may assume that all the necessary assumptions for the related test are satisfied by the data.
- (a) We will reject the null hypothesis at $\alpha = 0.01$
 - (b) We will reject the null hypothesis at $\alpha = 0.05$
 - (c) We will reject the null hypothesis at $\alpha = 0.1$
 - (d) We will reject the null hypothesis at $\alpha = 0.001$
 - (e) We fail to reject the null hypothesis at $\alpha = 0.1$
10. The weekly amount of money spent on maintenance and repairs by a company has a Normal distribution with mean \$400 and standard deviation \$ 20. If \$ 450 is budgeted for next week, what is the probability that the actual cost will exceed the budget? Choose the value closest to this probability.
- (a) 0.002
 - (b) 0.006
 - (c) 0.025
 - (d) 0.01
 - (e) 0.5
11. A random sample of size 4 is taken from a large (infinite) population with mean 100 and standard deviation 20 respectively. The mean and the standard deviation of the distribution of the sample mean are:
- (a) 100 and 20 respectively
 - (b) 50 and 10 respectively
 - (c) 50 and 20 respectively
 - (d) 100 and 10 respectively
 - (e) none of the above answers is correct

12. In a regression study, which of the following best describes the square of the correlation r^2 between the two variables?
- (a) The amount of error that is incurred from using a regression model to make predictions
 - (b) The amount of the variability explained by the standard deviations of the response and explanatory variables
 - (c) The proportion of the total variation in the response variable explained by the straight line relationship with the explanatory variable
 - (d) The sum of squares of the residuals
 - (e) None of the above
13. Which of the following pairs of events (denoted by A and B) are disjoint and independent?
- (a) Toss two coins, A = first toss is head, B = second toss is head
 - (b) Roll a die, A = observe an odd number, B = observe an even number
 - (c) Take two cards one by one from a deck of 52 cards, A = first card is red, B = both cards are red
 - (d) There are three kids in a family, A = two kids are girls, B = two kids are twins
 - (e) None of the above
14. A missile protection system consists of 3 radar sets operating independently, each with probability 0.9 of detecting a missile entering the zone that is covered by the system. If a missile enters the zone, what is the probability that at least one radar set will detect the missile?
- (a) 0.9
 - (b) 0.729
 - (c) 0.271
 - (d) 0.99
 - (e) 0.999

15. In a large mall a survey was taken. It was found that in a random sample of 45 women over the age of 25, 15 had children. Find the 90% confidence interval for the population proportion of women over the age of 25 in the mall who have children. (Choose the closest answer.)
- (a) (0.0208, 0.0653)
 - (b) (0.4337, 0.6774)
 - (c) (0.3333, 0.5555)
 - (d) (0.1374, 0.1626)
 - (e) (0.2177, 0.4489)
16. You have sampled 25 students to find the mean SAT scores. A 95% confidence interval for the mean SAT score is 900 to 1100. Which of the following statements gives a valid interpretation of this interval?
- (a) 95% of the 25 students have a mean score between 900 and 1100.
 - (b) 95% of the population of all students have a score between 900 and 1100.
 - (c) If this procedure were repeated many times, 95% of the sample means would be between 900 and 1100.
 - (d) If this procedure were repeated many times, 95% of the resulting confidence intervals would contain the true mean SAT score.
 - (e) If 100 samples were taken and a 95% confidence interval was computed, 5 of them would be in the interval from 900 to 1100.
17. A number calculated with complete population data and quantifies a characteristic of the population is called ...
- (a) Datum
 - (b) Statistics
 - (c) Population
 - (d) Stratum
 - (e) Parameter
18. In a group of 100 people, 40 own a cat, 25 own a dog, and 15 own a cat and a dog. Find the probability that a person chosen at random, owns a dog, given that he owns a cat.
- (a) 0.375
 - (b) 0.6
 - (c) 0.25
 - (d) 0.4
 - (e) 0.15

19. In clinical trials a certain drug has a 10% success rate of curing a known disease. If 15 people are known to have the disease. What is the probability of at least 2 being cured?
- (a) 0.5491
 - (b) 0.01
 - (c) 0.4509
 - (d) 0.15
 - (e) 0.7564
20. The exam scores in a class has a Normal distribution. 2.5 percent of the class failed (i.e. scored below 50) and 16 percent of the class got A's. A is the highest grade in this exam and a score of 86 or above qualifies for an A grade. Based on 68-95-99.7 % rule, what is the mean of this Normal distribution?
- (a) 56
 - (b) 62
 - (c) 68
 - (d) 74
 - (e) 78
21. In brief, what does the Central Limit Theorem say?
- (a) The area under a Normal density curve is one.
 - (b) Measures of central tendency should always be computed with and without outliers.
 - (c) For sufficiently large sample size, the sampling distribution of \bar{X} is approximately Normal.
 - (d) Confidence intervals have zero margin of error for large sample sizes.
 - (e) In the long run, the average outcome gets close to the distribution mean.
22. The events A and B are such that $P(A) = 2P(B)$, $P(A \text{ and } B) = 0.08$ and A and B are independent. Find $P(A \text{ or } B)$.
- (a) 0.68
 - (b) 0.6
 - (c) 0.4
 - (d) 0.52
 - (e) 0

23. In a certain game of chance, your chances of winning are 0.2 and it costs \$1 to play the game. If you win, you receive \$4 (for a net gain of \$3). If you lose, you receive nothing (for a net loss of \$1). You are going to play the game nine times. Assume that the outcomes are independent. Let T be the total net gain from this round of nine games. Find the standard deviation of T . Choose the closest.
- (a) \$ 1
 - (b) \$ 3
 - (c) \$ 5
 - (d) \$ 7
 - (e) \$ 9
24. The manufacturer of a bag of sweets claims that there is a 90% chance that the bag contains some toffees. If 20 bags are chosen, how many bags with toffees would you expect to have?
- (a) 20
 - (b) 2
 - (c) 9
 - (d) 18
 - (e) None
25. A statistics course has three tutorials, TUT01, TUT02 and TUT03, each with 50 students. TUT01 has 20 girls and 30 boys, TUT02 has 30 girls and 20 boys and TUT03 has 40 girls and 10 boys. We are going to select a sample of three students from this course by selecting one student at random from each tutorial. What is the probability that our sample will have at least two girls?
- (a) 0.288
 - (b) 0.464
 - (c) 0.656
 - (d) 0.784
 - (e) 0.944

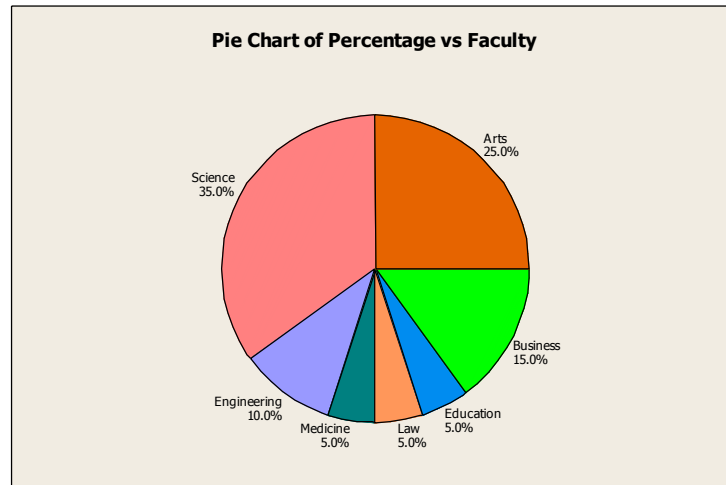
26. The distribution of bachelor's degrees conferred by a local college is listed below, by major.

Major	Frequency
English	2,073
Mathematics	2,164
Chemistry	318
Physics	856
Liberal Arts	1,358
Business	1,676
Engineering	868
Total	9,313

What is the probability that a randomly selected bachelor's degree is not in Mathematics?

- (a) 0.232
- (b) 0.768
- (c) 0.303
- (d) 0.682
- (e) 0.889

27. The pie chart below shows the percentage of students in each faculty at a university.



If the number of students in the faculty of Arts is 3000, then how MANY students are there in the faculty of Science?

- (a) 3400
 - (b) 3600
 - (c) 3800
 - (d) 4000
 - (e) 4200
28. Government data show that 10% of males under age 25 are unemployed. A random sample is taken of 400 males who are in the labor force and under age 25. Find the probability that the sample unemployment rate is 0.12 or more. (Hint: Use Normal approximation)
- (a) 0.1023
 - (b) 0.0918
 - (c) 0.9082
 - (d) 0.5470
 - (e) 0.4530

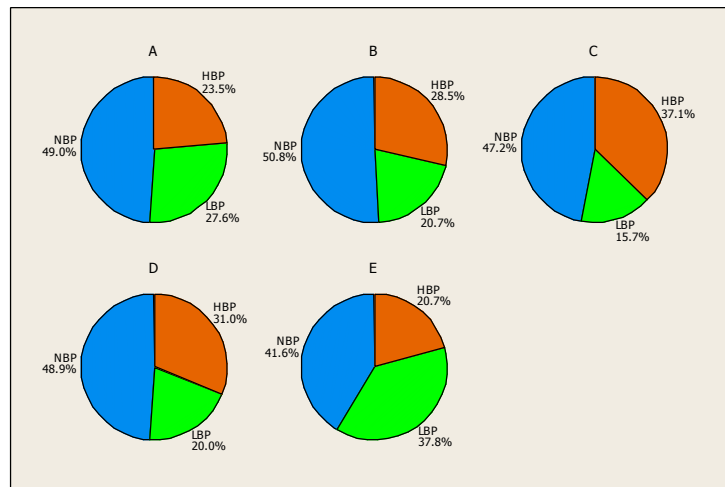
29. The unemployment rate (i.e. the percent unemployed in the labour force) in a certain city is 8%. A random sample of 150 people from the labour force in this city is drawn. Find the approximate probability that this sample contains fifteen or more unemployed people. Which of the following numbers is closest to this probability?

- (a) 0.1
- (b) 0.2
- (c) 0.3
- (d) 0.4
- (e) 0.5

30. In a study, a group of researchers measured the blood pressure levels of 474 patients, recording the blood pressure levels as high blood pressure (HBP), low blood pressure (LBP) or normal blood pressure (NBP). The patients were also classified into three age categories: under 30, 30-50, and over 50. The results of this study are summarized in the table below:

Blood pressure level	Under 30	30-50	Over 50	Total
HBP	23	51	73	147
LBP	27	37	31	95
NBP	48	91	93	232
Total	98	179	197	474

Which pie chart below represents the marginal distribution of blood pressure level?



- (a) pie chart A
- (b) pie chart B
- (c) pie chart C
- (d) pie chart D
- (e) pie chart E

31. A test of significance with $\alpha = 0.05$ results in a P-value of 0.1. What do you conclude?
- (a) Reject the null hypothesis.
 - (b) Reject the alternative hypothesis.
 - (c) Do not reject the alternative hypothesis.
 - (d) Do not reject the null hypothesis.
 - (e) Accept the null hypothesis.
32. There are ten children in a room. The mean and standard deviation of their ages are 5 years and 1.25 years respectively. If another 5-year-old child enters this room, what will happen to the mean and standard deviation of the ages of the children in the room?
- (a) The mean will stay the same but the standard deviation will increase.
 - (b) The mean and standard deviation will both increase.
 - (c) The mean and standard deviation will both decrease.
 - (d) The mean will stay the same but the standard deviation will decrease.
 - (e) The mean and standard deviation will both stay the same.
33. Two stores (store A and store B) sell watermelons. The weights of watermelons at store A has a Normal distribution with mean 21 pounds and standard deviation 3 pounds. The weights of watermelons at store B has a Normal distribution with mean 18 pounds and standard deviation 2 pounds. The weights of watermelons at the two stores are independent. If you select a melon at random at each store, what is the probability that melon you selected at store B will be heavier? Choose the closest.
- (a) 0.1
 - (b) 0.2
 - (c) 0.3
 - (d) 0.4
 - (e) 0.5
34. The regression equation relating midterm score (x) and final exam score (y) in a course is $y = 20 + 0.8x$. Which of the following is a valid conclusion based on this information?
- (a) If Adam has scored 10 points higher than Bob on the midterm, then Adam will score 8 points higher than Bob on the final exam.
 - (b) Students who score 80 on the midterm are predicted to score 84 on the final.
 - (c) If a student scores 0 on the midterm, then that student will score 20 on the final.
 - (d) On average, the students' scores on the final exam will only be 80% of what they were on the midterm.
 - (e) This equation shows that the midterm and the final exam scores are not related.

35. The heights of women in a certain population have a Normal distribution with mean 64.5 inches and standard deviation 2.5 inches. We select three women at random from this population. Assume that their heights are independent. Find the probability that the tallest of these three women will be taller than 67 inches. Which of the following numbers is closest to this probability?
- (a) 0.0
 - (b) 0.2
 - (c) 0.4
 - (d) 0.6
 - (e) 0.8
36. X and Y are two independent random variables and $Var(X) = 0.61$ and $Var(Y) = 0.21$. Find $SD(X - Y)$ (i.e. σ_{X-Y}).
- (a) 0.6451
 - (b) 0.6324
 - (c) 0.9055
 - (d) 0.5727
 - (e) 0.3227
37. A consumer product agency tests mileage (miles per gallon) for a sample of automobiles using each of four different types of gasoline. Which of the following statements regarding this study is true?
- (a) There are four explanatory variables and one response variable in this study.
 - (b) There is only one explanatory variable with four levels and one response variable in this study.
 - (c) Miles per gallon is the only explanatory variable in this study.
 - (d) This study has a single response variable with four levels.
 - (e) None of the above statements is true.

38. The two-way table below classifies the members of a fitness club by gender and food habits. Use this information to answer this question and the next question.

	Women	Men
Vegetarian	9	3
Non-vegetarian	8	10

If a member is selected at random from the club, what is the probability that this person is a vegetarian female?

- (a) $3/4$
 - (b) $3/10$
 - (c) $1/10$
 - (d) $4/9$
 - (e) $9/17$
39. Using the information in question 38 above, what is the probability that a randomly selected member is non-vegetarian?
- (a) $2/5$
 - (b) $1/3$
 - (c) $4/15$
 - (d) $3/5$
 - (e) $8/17$
40. It is widely thought that there is a high incidence of disability among the homeless population. A random sample of 110 homeless people contained 84 who were disabled on one or more categories (such as psychiatric disability, medical disability etc.). Let p denote the proportion of all homeless people having one or more types of disability. Which of the following numbers is the closest to the value of the test Z-statistic for testing the null hypothesis $H_0 : p = 0.75$ against $H_a : p \neq 0.75$?
- (a) 0.0
 - (b) 0.3
 - (c) 0.6
 - (d) 0.9
 - (e) 1.2

41. Two independent random variables have the following distributions:

x	-1	0	1
p	0.2	0.3	0.5

y	1	2
p	0.3	0.7

Find the mean of $Y - 2X$ (i.e. find $E(Y - 2X)$).

- (a) 1.1
- (b) -3.1
- (c) 0.3
- (d) 1.7
- (e) 1.4

42. The regression equation and the summary statistics given below were obtained from a regression analysis of the relation between murder rate (y , the number of murders per 1,000,000 inhabitants per annum) and unemployment (x , the percentage of the labour force unemployed) using data for 20 cities.

Descriptive Statistics: x , y

Variable	N	Mean	StDev	Minimum	Q1	Q3	Maximum
x	20	deleted	1.207	4.900	6.050	8.125	9.300
y	20	20.57	9.88	5.30	12.88	26.63	40.70

Regression Analysis: y versus x

The regression equation is $y = - 28.5 + 7.08 x$

What is the correlation between x and y ?

- (a) 0.035
- (b) 0.0708
- (c) 0.122
- (d) 0.865
- (e) 0.988

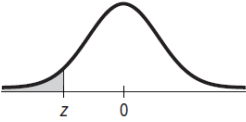
43. Using the information in question 42 above, calculate the average unemployment rate (i.e. the value of \bar{x}) of these 20 cities. Note: This value has been deleted in the summary statistics given in question 42 above but can be calculated from other information given.

Which of the following numbers is closest to this value?

- (a) 7
- (b) 14
- (c) 21
- (d) 28
- (e) 49

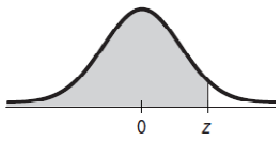
44. One of the cities in the data set used in question 42 above, had unemployment rate 6.0 (i.e. $x = 6.0$) and murder rate 14.5. What is the residual for this city?

- (a) 0.52
- (b) -68.16
- (c) -27.98
- (d) 43
- (e) -96.66

<p>Table Z Areas under the standard Normal curve</p> 		Second decimal place in z									z	
		0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01		0.00
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0000 [†]	-3.9
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	-3.8
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	-3.7
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	-3.6
0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	-3.5
0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	-3.4
0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005	0.0005	-3.3
0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007	0.0007	-3.2
0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0009	0.0010	-3.1
0.0010	0.0010	0.0011	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013	0.0013	-3.0
0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0016	0.0017	0.0018	0.0018	0.0018	0.0019	-2.9
0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0023	0.0024	0.0025	0.0025	0.0026	-2.8
0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0031	0.0032	0.0033	0.0034	0.0034	0.0035	-2.7
0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0041	0.0043	0.0044	0.0045	0.0045	0.0047	-2.6
0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0055	0.0057	0.0059	0.0060	0.0060	0.0062	-2.5
0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0073	0.0075	0.0078	0.0080	0.0080	0.0082	-2.4
0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0096	0.0099	0.0102	0.0104	0.0104	0.0107	-2.3
0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0125	0.0129	0.0132	0.0136	0.0136	0.0139	-2.2
0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0162	0.0166	0.0170	0.0174	0.0174	0.0179	-2.1
0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0207	0.0212	0.0217	0.0222	0.0222	0.0228	-2.0
0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0262	0.0268	0.0274	0.0281	0.0281	0.0287	-1.9
0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0329	0.0336	0.0344	0.0351	0.0351	0.0359	-1.8
0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0409	0.0418	0.0427	0.0436	0.0436	0.0446	-1.7
0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0505	0.0516	0.0526	0.0537	0.0537	0.0548	-1.6
0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0618	0.0630	0.0643	0.0655	0.0655	0.0668	-1.5
0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0749	0.0764	0.0778	0.0793	0.0793	0.0808	-1.4
0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0901	0.0918	0.0934	0.0951	0.0951	0.0968	-1.3
0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1075	0.1093	0.1112	0.1131	0.1131	0.1151	-1.2
0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1271	0.1292	0.1314	0.1335	0.1335	0.1357	-1.1
0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1492	0.1515	0.1539	0.1562	0.1562	0.1587	-1.0
0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1736	0.1762	0.1788	0.1814	0.1814	0.1841	-0.9
0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2005	0.2033	0.2061	0.2090	0.2090	0.2119	-0.8
0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2296	0.2327	0.2358	0.2389	0.2389	0.2420	-0.7
0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2611	0.2643	0.2676	0.2709	0.2709	0.2743	-0.6
0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2946	0.2981	0.3015	0.3050	0.3050	0.3085	-0.5
0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3300	0.3336	0.3372	0.3409	0.3409	0.3446	-0.4
0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3669	0.3707	0.3745	0.3783	0.3783	0.3821	-0.3
0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4052	0.4090	0.4129	0.4168	0.4168	0.4207	-0.2
0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4443	0.4483	0.4522	0.4562	0.4562	0.4602	-0.1
0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4840	0.4880	0.4920	0.4960	0.4960	0.5000	-0.0

[†]For $z \leq -3.90$, the areas are 0.0000 to four decimal places.

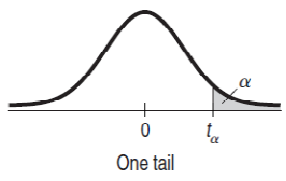
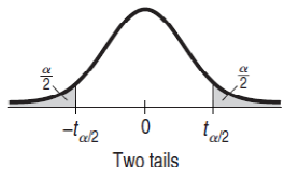
Table Z (cont.)
Areas under the
standard Normal curve



z	Second decimal place in z									
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000 [†]									

[†]For $z \geq 3.90$, the areas are 1.0000 to four decimal places.

Two-tail probability	0.20	0.10	0.05	0.02	0.01		
One-tail probability	0.10	0.05	0.025	0.01	0.005		
Table T	df					df	
Values of t_α	1	3.078	6.314	12.706	31.821	63.657	1
	2	1.886	2.920	4.303	6.965	9.925	2
	3	1.638	2.353	3.182	4.541	5.841	3
	4	1.533	2.132	2.776	3.747	4.604	4
	5	1.476	2.015	2.571	3.365	4.032	5
	6	1.440	1.943	2.447	3.143	3.707	6
	7	1.415	1.895	2.365	2.998	3.499	7
	8	1.397	1.860	2.306	2.896	3.355	8
	9	1.383	1.833	2.262	2.821	3.250	9
	10	1.372	1.812	2.228	2.764	3.169	10
	11	1.363	1.796	2.201	2.718	3.106	11
	12	1.356	1.782	2.179	2.681	3.055	12
	13	1.350	1.771	2.160	2.650	3.012	13
	14	1.345	1.761	2.145	2.624	2.977	14
	15	1.341	1.753	2.131	2.602	2.947	15
	16	1.337	1.746	2.120	2.583	2.921	16
	17	1.333	1.740	2.110	2.567	2.898	17
	18	1.330	1.734	2.101	2.552	2.878	18
	19	1.328	1.729	2.093	2.539	2.861	19
	20	1.325	1.725	2.086	2.528	2.845	20
	21	1.323	1.721	2.080	2.518	2.831	21
	22	1.321	1.717	2.074	2.508	2.819	22
	23	1.319	1.714	2.069	2.500	2.807	23
	24	1.318	1.711	2.064	2.492	2.797	24
	25	1.316	1.708	2.060	2.485	2.787	25
	26	1.315	1.706	2.056	2.479	2.779	26
	27	1.314	1.703	2.052	2.473	2.771	27
	28	1.313	1.701	2.048	2.467	2.763	28
	29	1.311	1.699	2.045	2.462	2.756	29
	30	1.310	1.697	2.042	2.457	2.750	30
	32	1.309	1.694	2.037	2.449	2.738	32
	35	1.306	1.690	2.030	2.438	2.725	35
	40	1.303	1.684	2.021	2.423	2.704	40
	45	1.301	1.679	2.014	2.412	2.690	45
	50	1.299	1.676	2.009	2.403	2.678	50
	60	1.296	1.671	2.000	2.390	2.660	60
	75	1.293	1.665	1.992	2.377	2.643	75
	100	1.290	1.660	1.984	2.364	2.626	100
	120	1.289	1.658	1.980	2.358	2.617	120
	140	1.288	1.656	1.977	2.353	2.611	140
	180	1.286	1.653	1.973	2.347	2.603	180
	250	1.285	1.651	1.969	2.341	2.596	250
	400	1.284	1.649	1.966	2.336	2.588	400
	1000	1.282	1.646	1.962	2.330	2.581	1000
	∞	1.282	1.645	1.960	2.326	2.576	∞
Confidence levels	80%	90%	95%	98%	99%		



Binomial Distribution Table

n	k	0.1	0.2	0.3	0.4	0.5	0.6
14	0	0.2288	0.0440	0.0068	8e-04	1e-04	0.0000
14	1	0.3559	0.1539	0.0407	0.0073	9e-04	1e-04
14	2	0.2570	0.2501	0.1134	0.0317	0.0056	5e-04
14	3	0.1142	0.2501	0.1943	0.0845	0.0222	0.0033
14	4	0.0349	0.1720	0.2290	0.1549	0.0611	0.0136
14	5	0.0078	0.0860	0.1963	0.2066	0.1222	0.0408
14	6	0.0013	0.0322	0.1262	0.2066	0.1833	0.0918
14	7	2e-04	0.0092	0.0618	0.1574	0.2095	0.1574
14	8	0.0000	0.0020	0.0232	0.0918	0.1833	0.2066
14	9	0.0000	3e-04	0.0066	0.0408	0.1222	0.2066
14	10	0.0000	0.0000	0.0014	0.0136	0.0611	0.1549
14	11	0.0000	0.0000	2e-04	0.0033	0.0222	0.0845
14	12	0.0000	0.0000	0.0000	5e-04	0.0056	0.0317
14	13	0.0000	0.0000	0.0000	1e-04	9e-04	0.0073
14	14	0.0000	0.0000	0.0000	0.0000	1e-04	8e-04

n	k	0.1	0.2	0.3	0.4	0.5	0.6
15	0	0.2059	0.0352	0.0047	5e-04	0.0000	0.0000
15	1	0.3432	0.1319	0.0305	0.0047	5e-04	0.0000
15	2	0.2669	0.2309	0.0916	0.0219	0.0032	3e-04
15	3	0.1285	0.2501	0.1700	0.0634	0.0139	0.0016
15	4	0.0428	0.1876	0.2186	0.1268	0.0417	0.0074
15	5	0.0105	0.1032	0.2061	0.1859	0.0916	0.0245
15	6	0.0019	0.0430	0.1472	0.2066	0.1527	0.0612
15	7	3e-04	0.0138	0.0811	0.1771	0.1964	0.1181
15	8	0.0000	0.0035	0.0348	0.1181	0.1964	0.1771
15	9	0.0000	7e-04	0.0116	0.0612	0.1527	0.2066
15	10	0.0000	1e-04	0.0030	0.0245	0.0916	0.1859
15	11	0.0000	0.0000	6e-04	0.0074	0.0417	0.1268
15	12	0.0000	0.0000	1e-04	0.0016	0.0139	0.0634
15	13	0.0000	0.0000	0.0000	3e-04	0.0032	0.0219
15	14	0.0000	0.0000	0.0000	0.0000	5e-04	0.0047
15	15	0.0000	0.0000	0.0000	0.0000	0.0000	5e-04