Discrete Random Variables

• A random variable with the set of possible values at most countable is called **discrete**.

• Say that a discrete random variable $X : S \to \{a_1, a_2, \ldots, a_n, \ldots\}$ and define

$$p : \{a_1, a_2, \ldots, a_n, \ldots\} \to [0, 1]$$

as $p(a_i) = P(X = a_i)$. The function $p(\cdot)$ is called the *probability mass function*.

• The randomness of X is perfectly described by its probability mass function.

• A probability mass function $p(\cdot)$ must satisfy

1. $p(a_i) \geq 0$, for all $i \geq 0$.

2. For any $x \notin \{a_1, a_2, \ldots, a_n, \ldots\}$, $p(x) = 0$,

3. $\sum_{i=1}^{\infty} p(a_i) = 1$.

$\boxed{\textbf{Mean of discrete random variables}}$

• Let $X : S \to \{a_1, a_2, \ldots, a_n, \ldots\}$ be a discrete random variable with probability mass function $p(\cdot)$. The mean (a.k.a expected value) of the random variable X is denoted $E[X]$ ("E" stands for "expected value") and is defined as

$$E[X] = \sum_{k=1}^{\infty} a_k \cdot p(a_k).$$

Example The expected value of $X \sim Po(\lambda)$ is

$$E[X] = \sum_{k=0}^{\infty} k \cdot p(k) = \lambda.$$

Properties of $E[X]$

1. For any X and any real number c,

$$E[c + X] = c + E[X].$$

2. For any X and any real number c,

$$E[c \cdot X] = c \cdot E[X].$$

3. For any real number c, $E[c] = c$.

$\boxed{\textbf{Properties of the Expectation -cont'd}}$

• **"The Law of the Unconscious Probabilist"**: If X is a random variable $X : S \to \{a_1, ..., a_n, ...\}$ with probability mass function $p(\cdot)$ and g is a map $g : \{a_1, ..., a_n, ...\} \to R$, then

$$E[g(X)] = \sum_{i=1}^{\infty} g(a_i)p(a_i)$$

• The physical interpretation of the expected value of a probability distribution as the coordinate of the center of gravity.

Variance of a Random variable

- The variance is a measure of the "spread" of a distribution.

- The variance of a random variable X is

$$Var(X) = E[(X - E[X])^2]$$

- Alternative form: $Var(X) = E[X^2] - (E[X])^2$

Properties of the variance

1. $Var(c + X) = Var(X)$ for any constant c.

2. $Var(cX) = c^2 Var(X)$ for any constant c.

$\boxed{\textbf{The Bernoulli distribution}}$

• Suppose a trial is performed such that its outcome can be classified as either a "success" or as a "failure". Such a trial is called a *Bernoulli trial*.

• Define the random variable $X : S \to \{0, 1\}$ as being equal to 1 if the outcome is a success and equal to 0 if the outcome is a failure.

• If the probability of a success is p, then the probability mass function is

$$p(1) = p, p(0) = 1 - p.$$

The Binomial distribution

• Suppose that n *independent* Bernoulli trials are performed and each of them has *the same probability of success* equal to p. If $X : S \to \{0, 1, \dots, n-1, n\}$ is the random variable equal to the number of successes in the n trials then X is said to follow the *binomial distribution with parameters* n *and* p and is denoted $X \sim Bin(n, p)$.

• The probability mass function of X is given by

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k}, \forall k \in \{0, 1, \dots, n\}.$$

• If $X \sim Bin(n, p)$ then as k goes from 0 to n, $p(k)$ first increases monotonically and then decreases monotonically reaching its largest value when k is the largest integer less than or equal to $(n+1)p$.

• If for each $0 \leq i \leq n$, Y_i is the Bernoulli random variable associated to the i-th Bernoulli trial, that is, $Y_i = 1$ if the i-th trial was a success and $Y_i = 0$ otherwise, then

$$X = \sum_{i=1}^{n} Y_i.$$

The Poisson distribution and Poisson Process

• A random variable $X : S \to \{1, 2, \ldots, n, \ldots\}$ is said to follow the Poisson distribution with parameter $\lambda > 0$ and is denoted $X \sim Po(\lambda)$ if for any integer $i \geq 0$

$$p(i) = e^{-\lambda}\frac{\lambda^i}{i!}.$$

• Let $N(t)$ denote the number of events occuring by time t. Assume that the following hold:

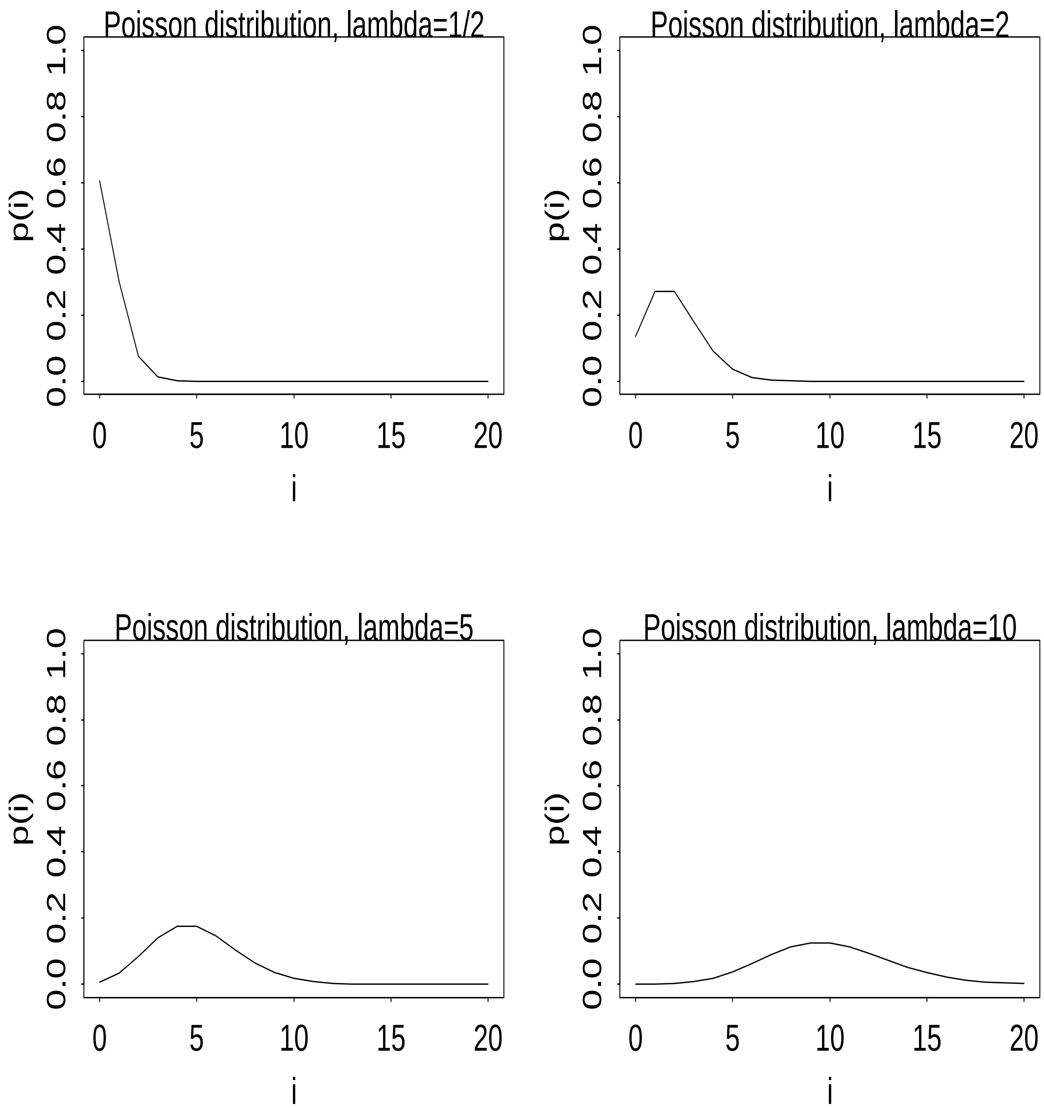
- stationarity - for two time intervals of equal length the distribution of the number of events is the same within each interval

- independent increment - the number of occurences in an interval $(t, t + s)$ does not depend on the number of occurences from previous times.

- orderliness - no two (or more) events can occur simultaneously.

- $N(0) = 0$ - there is a time origin at which the counting of events starts.

Then there exists a positive number $\lambda > 0$ such that the distribution of $N(t)$ is $Poisson(\lambda t)$ for all $t > 0$.



Poisson distribution, lambda=1/2
Poisson distribution, lambda=2
Poisson distribution, lambda=5
Poisson distribution, lambda=10
p(i)
0.0 0.2 0.4 0.6 0.8 1.0
0 5 10 15 20
i

The Geometric distribution

• Suppose that a sequence of independent Bernoulli trials each with probability of success p are performed.

• Let X be the number of experiments until the first success occurs. X is a *geometric random variable*. The probability mass function for X is $p(X = n) = (1 - p)^{n-1}p$ for all $n > 0$.

• The distribution of X is called geometric. Notation $Geo(p)$

Negative Binomial distribution

- Generalizes the geometric distribution.

- Suppose that a sequence of independent Bernoulli trials each with probability of success p are performed.

- Let X be the number of experiments until the first r successes occur. X is a *a negative binomial random variable*. The probability mass function for X is $p(X = n) = \binom{n-1}{r-1}(1-p)^{n-r}p^r$ for all $n \geq r$. Notation $NBin(r, p)$

Table of means and variances

Distribution	Mean	Variance
Bernoulli(p)	p	p(1-p)
Bin(n, p)	np	$np(1-p)$
Poisson(λ)	λ	λ
Geo(p)	$\frac{1}{p}$	$\frac{1-p}{p^2}$
NBin(r, p)	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$