binomial

iid Bernoulli trials yielding

\[ X_1 \sim \text{iid } X, \quad X_2 \sim \text{iid } X, \ldots, \quad X_N \sim \text{iid } X \]

Note possible values of \( X \) are \( \binom{0}{0} \) or \( \binom{1}{0} \)

The pdf of the components of \( X \) is just the

The pdf of \( X \)

\[ G_X(z) = E(z^X) = E(z_1^{1st \text{ comp of } X} z_2^{2nd \text{ comp of } X}) \]

\[ = z_1 p_1 + z_2 p_2 \]

\[ = p_1 z_1 + p_2 z_2 \]

= pgf of a "vector Bernoulli"

Now set

\[ Y = X_1 + \ldots + X_N \]

\[ \Rightarrow G_Y(z) = (p_1 z_1 + p_2 z_2)^N = \sum_{y_1, y_2 = 1}^N (y_1, y_2)^N p_1^{y_1} p_2^{y_2} z_1^{y_1} z_2^{y_2} \]

\[ \Rightarrow p(Y = y) = \binom{N}{y_1, y_2} p_1^{y_1} p_2^{y_2} \]
Note \( z_2 = 1 \Rightarrow (p, z_1 + z_2)^N \) is the pgf of the 1st component of \( X \) (this is the binomial\((N, p)\) as you knew it).

Extend to the multinomial

\[
\begin{align*}
X_1, X_2, \ldots, X_M & \overset{iid}{\sim} X \\
Mx1 & \quad Mx1
\end{align*}
\]

\( X \) has one component \( = 1 \) + the rest \( 0 \). The probability that the \( 1 \) is in the \( i \)th place we call \( p_i \) \((i=1, \ldots, M)\). The pgf \( G \) of \( X \) is

\[
(p, z_1 + z_2 + \cdots + z_M)
\]

Letting

\[
Y = X_1 + \cdots + X_M
\]

we get

\[
G(Y) = (p, z_1 + \cdots + z_M)^N
\]

\[
\Rightarrow \quad P(Y = y) = \binom{N}{y_1, \ldots, y_M} p_1^{y_1} \cdots p_M^{y_M}
\]

\[
\binom{N}{y_1, \ldots, y_M}
\]
\[ \frac{d^2}{dz_1 dz_2} G(z) = E(Y, z_1^{-1} z_2 z_2^{-1} z_3^{-1} \ldots) \] works for all counting vectors.

Can use to get the \( \text{cov}(X_i, Y_j) \) for counting vectors \( Y \).

**Properties of covariances**

\[ \text{cov}(X, Y) = E(XY) - E(X)E(Y) \]

\[
\begin{align*}
\text{cov}(X, X) &= \text{Var}(X) \\
\text{cov}(X+c, Y+d) &= \text{cov}(X, Y) \\
\text{cov}(aX, bY) &= ab \text{cov}(X, Y) \\
\text{cov}(\sum_{i} X_i, \sum_{j} Y_j) &= \sum_{i,j} \text{cov}(X_i, Y_j)
\end{align*}
\]

**eg** \( X_1 \sim \text{Poisson}(\lambda_1), X_2 \sim \text{Poisson}(\lambda_2), X_3 \sim \text{Poisson}(\lambda_3) \)

\[ U = X_1 + X_2 \quad \text{easy to get the pdf} \]

\[ V = X_2 + X_3 \]

\[ E(z_1^r z_2^s) \]

\[ = E(z_1^{X_1+X_2} z_2^{X_2+X_3}) \]
\[ E \left( z_1 X_1 \left( z_2 z_2 \right)^2 z_2 \right) \]
\[ = E \left( z_1 X_1 \right) E \left( z_2 z_2 \right)^2 E \left( z_2 \right) \]
\[ = e \left( z_1 \right) e \left( z_2 \right)^2 e \left( z_2 \right) \]
\[ = e \left( z_1 \right) e \left( z_2 \right)^2 e \left( z_2 \right) \]

- Use this to get \( E(2UV) \) — imp \( \}
  15 minutes

and then get \( \text{cov}(U,V) \)

Another way
\[
\text{cov}(U,V) = \text{cov}(X_1 + X_2, X_2 + X_3) \\
= \text{cov}(X_1, X_2) + \text{cov}(X_1, X_3) + \text{cov}(X_2, X_2) + \text{cov}(X_2, X_3) \\
= \text{cov}(X_2, X_2) = \text{Var}(X_2) = \lambda_2
\]

Note: \( X, Y \) ind \( \Rightarrow \) \( E(XY) = E(X)E(Y) \) \( \Rightarrow \text{cov}(X,Y) = 0 \)

\( \exists Z \sim N(0,1) \). Let \( X = Z \), \( Y = Z^2 \). Then \( X \& Y \) are dependent

\[
E(XY) = E(ZZ^2) = E(Z^3) = 0
\]

\( \therefore X \& Y \) are uncorrelated.
Spatial Poisson process

Background. A rv is uniform on a set if its PDF is constant there.

Throw N points onto V in a uniform way.

everything oriented $A_1, \ldots, A_{M-1}$

$\{A_1, \ldots, A_{M-1}, A_{M}\}$ is a partition of $V$.

Let $N(A) = \#$ of pts in a set $A \subset V$.

$\sim$ binomial $(N, |A|/|V|)$

Also $N(A_1), \ldots, N(A_M)$ is the multinomial with $p_i = |A_i|/|V|$

The PDF is

$$(p_1^Z, \ldots, p_{M-1}^Z_{M-1}, p_M^Z_{M})^N$$
In the limit you get a Poisson point process on $\mathbb{R}^d$ of "rate" $\rho$.

Poisson spatial process, Poisson counting process, Poisson process.

Why is it called Poisson?
\[(*) \quad \left[ P_1(z_{i-1}) + \cdots + P_{M-1}(z_{M-1}) + 1 \right]_N \]

\[ P_i = \frac{|A_i|}{|V|} = \frac{|A_i|}{N} \cdot \frac{N}{|V|} = \frac{\rho |A_i|}{N} \]

so that \((*)\) is

\[ \left[ 1 + \frac{\rho |A_1|(z_{i-1}) + \rho |A_2|(z_{i-1}) + \cdots + \rho |A_{M-1}|(z_{M-1})}{N} \right]^N \]

\[
\lim_{N \to \infty} \rho |A_1|(z_{i-1}) + \cdots + \rho |A_{M-1}|(z_{M-1})
\]

\[ \Rightarrow \mathcal{C} \]

\[ = \text{PDF of } M-1 \text{ independent Poisson } \]

\[ \text{with means } \rho |A_1|, \ldots, \rho |A_{M-1}| \]

**Note:** For a Poisson process, the \#'s in disjoint regions are independent Poisson, \( \lambda + \) the \# in a set \( A \) is Poisson with mean \( \frac{\rho \cdot \lambda}{\text{usual}} \)
Special case: Poisson process of rate $\lambda$ on $t \geq 0$.

$$\begin{array}{c}
| 0 | t_1 | t_2 | t_3 | \ldots | \text{time} \\
\hline
\end{array}$$

$\forall t_0 < t_1 < t_2 < \ldots \quad N(t) = \# \text{ of points in } [0, t]$.

Then $N(t_i) - N(t_{i-1}) = \# \text{ of pts in } (t_{i-1}, t_i]$, $i = 1, 2, \ldots$.

$\{N(t) : t \geq 0\}$ is a Poisson counting process of rate $\lambda$ on $t \geq 0$.

Note that the $N(t_i) - N(t_{i-1})$ are independent (independent increments).

Times between pts are \textit{i.i.d.} \{\text{want to study}\}

Time to the $n$th pt is a \textit{rv}
Let \( S_n \) = time to the \( n \)th pt. +
\( X_1, X_2, \ldots \) be the times between pts
\( \uparrow \) time from 1st pt to 2nd
\( \uparrow \) time from 0 to the 1st pt

dist'n of \( X_1 \)

Let \( F(x) = P(X_1 \leq x) \) = def of \( X_1 \)
\( F(x) = P(X_1 > x) \) = Tail probability

Assume \( F \) (or \( 1 - F \)) determine the dist'n
(which it does but hard to prove).

Here
\( F(x) = 0 \) unless \( x > 0 \)

Let \( x > 0 \). Then
\( 1 - F(x) = P(X_1 > x) = P(N(x) = 0) \)

But \( N(x) \sim \text{Poisson}(\lambda x) \)

\( \Rightarrow P(N(x) = 0) = e^{-\lambda x} \)
\[ F(x) = e^{-\lambda x}, \quad x > 0 \]

\[ F(x) = 1, \quad x \leq 0 \]

Notice \( F'(x) = -\lambda e^{-\lambda x}, \quad x > 0 \)

\[ F'(x) = \lambda e^{-\lambda x}, \quad x > 0 \]

which is the exponential (1) pdf. So \( X \sim \text{exponential}(\lambda) \)

\[ \text{N/t. If a pdf exists then } F'(x) = -\frac{1}{f(x)}. \]

The exponential(1) distr'n, by def'n, has pdf

\[ f(x) = \lambda e^{-\lambda x}, \quad x > 0 \]

\[ \text{mean} = \frac{1}{\lambda} \]

\[ = 0, \quad \text{otherwise} \]

If \( F'(x) \) is a pdf then we are dealing with an absolutely cont' a distr'n.

\[ \text{For } S_n = \text{time to } n \text{th point} \]

\[ \overline{F}(x) = P(S_n > x) \]
Let $x > 0$ then
\[
\bar{F}_n(x) = P(S_n > x) = P(N(x) \leq n - 1)
\]
\[
= \sum_{k=0}^{n-1} e^{-\lambda x} \frac{(\lambda x)^k}{k!}
\]

Now calculate $\bar{F}_n'(x)$ to get the pdf of a gamma (check this).

---

Extra material (basic conditioning, a review)

Def'm if $P(A) > 0$ define $E(Y|A) = \frac{E[XY|A]}{E[I(A)]]}$

It is easily seen that, for fixed $A$, $E(\cdot | I(A))$ satisfies the Axioms for $E$ (verify this).

Def'm $P(B|A) = E(I(B)|A)$

Of course $P(B|A) = P(AB)/P(A)$ as before
If $X$ is discrete then we set
\[ r(x) = E(Y | X = x) \]
and
\[ E(Y | X) = r(X) \]
If $Y$ is also discrete then it is easily verified
\[ E(Y) = E[E(Y | X)] \]
and
\[ E(Y | X_1) = E[E(Y | X_1, X_2) | X_1] \]
A further property is
\[ E[Y g(X)] = E[E(Y | X) g(X)] \]
"A" real $g$.

These last three properties are true in general and are easily verified when $(\tilde{Y})$
is its with $pdf \ \tilde{f}(y, x').$

Finally, it can be shown that $E(Y | X)$
satisfies (see Th 5.3.1)
\[ E(Y - E(Y | X))^2 = \min_{h} E(Y - h(X))^2 \]
Application of the uniform + Bernoulli / Indicator

N - towns ≥ at most 1 road between m - roads ≥ any 2 towns

\[ \frac{2}{3} \text{ of roads go into to dotted collection} \]

In fact there is always at collection of towns ≥ at least half of roads lead into it.

Sol'n Go to each town + toss a fair coin. If H put town into a collection S. S is a random collection of towns. Note that there are \( 2^N \) possible "values" that S can be. Let \( X = \# \) of roads leading into S from outside. We need to show there is
a possible value of \( X \geq m/2 \). Since \( X \) is a (positive) counting rv this will hold if \( E(X) > m/2 \). Now, let

\[ I_j = I(\{ \text{road } j \text{ leads into } S \}) \]

Then

\[ X = \sum_{j=1}^{m} I_j \]

and

\[ E(X) = m \cdot E(I_i) = m \cdot P(\text{road } i \text{ leads into } S) \]

\[ = m \cdot P(\text{one } H \text{ and one } T \text{ for the } 2 \text{ tosses}) \]

\[ = m/2 \]

Let \( \{ N(t) : t > 0 \} \) be a Poisson process of rate \( \lambda \) on \( t > 0 \).

Suppose you know \( N(t_0) = m \). Call the times of the points \( T_1 < T_2 < \ldots < T_m \)
We want the pdf of $T_r$.

**Approach**

Use def $F$ so $\overline{F} = 1 - F$

**Aside**

$F(x) = P(X \leq x), \forall x \in \mathbb{R}$

For a cts rv $X$

$$F(x) = \int_{-\infty}^{x} f(t) \, dt$$

$\Rightarrow$ at continuity pts

$$F'(x) = f(x)$$

In fact we could take $F'$ to be an equivalent pdf.

$$F(b) - F(a) = P(a < X \leq b)$$

$$\Rightarrow \quad \underbrace{P(X \leq a)}_{F(a)} + P(a < X \leq b) = P(X \leq b)$$

$$= \int_{a}^{b} F'(t) \, dt \quad [\text{FLAC}]$$
Remark, if \( F \) is cts it is possible \( F' = f \). For \( f \) it is almost every \( x \) for \( F'(x) = 0 \) for almost every \( x \) (places where not true have length 0).

So in that case \( \int_{-\infty}^{\infty} F'(x) \, dx = 0 \).

\( F'(x) = -f(x) \). So knowing \( F \) \( \Rightarrow \) knowing the distribution.

By the way,
\( x_1 < x_2 \Rightarrow \mathbb{P}(X \leq x_1) \leq \mathbb{P}(X \leq x_2) \)
\( \Rightarrow \{ X \leq x_1 \} \Rightarrow \{ X \leq x_2 \} \)

\( \Rightarrow F \) is increasing.
\( \forall x_n \downarrow a \Rightarrow \{ X \leq x_n \} \downarrow \{ X \leq a \} \)
\( \Rightarrow \mathbb{P}(X \leq x_n) \rightarrow \mathbb{P}(X \leq a) \)
(continuity property of \( \mathbb{P} \))

\( \Rightarrow F \) is right cts.
\[ x \nearrow \infty \Rightarrow \{X \leq x\} \uparrow \{X < \infty\} \]
\[ \Rightarrow F(x_m) \to P(X < \infty) = 1 \]

\[ x \searrow \infty \Rightarrow F(x) \to 1 \]
\[ F(\infty) = 1 \]

\[ x \downarrow -\infty \Rightarrow \exists \{X \leq x\} \downarrow \emptyset \]
\[ \Rightarrow F(x_m) \to 0 \]

\[ x \searrow -\infty \Rightarrow F(x) \to 0 \]
\[ F(-\infty) = 0 \]

Notation:
\[ x_m \uparrow a \Rightarrow \{X \leq x_m\} \uparrow \{X < a\} \]

\[ \lim_{x \uparrow a} F(x) = P(X < a) \neq F(a) \]

\[ P(X = a) = F(a) - \lim_{x \uparrow a} F(x) \]

Knowing \( F \) implies knowing \( P(a \leq X < b) \)

Knowing the distribution function \( F \) implies knowing the distribution function \( F \).
Back to our Poisson process.

\[ F_{(m)}(x) = P(T_m \leq x) \]

\[ = P(T_m \leq x \mid N(t_o) = m) \]

\[ = P(N((x, t_o]) = 0 \mid N(t_o) = m) \]

\[ = P(N((x, t_o]) = 0, N(t_o) = m) \]

\[ \frac{P(N(t_o) = m)}{P(N(t_o) = m)} \]

\[ = P(N(x) = m, N((x, t_o]) = 0) \]

\[ \frac{P(N(t_o) = m)}{P(N(t_o) = m)} \]

\[ = \frac{P(N(x) = m) P(N((x, t_o]) = 0)}{P(N(t_o) = m)} \]
\[ = P(N(x) = n) \cdot P(N((x, \infty]) = 0) \]
\[ P(N(t_0) = m) \]
\[ = e^{-\lambda x} \left( \frac{dx}{m!} \right)^m \cdot e^{-\lambda (t_0 - x)} \]
\[ e^{-\lambda t_0} \cdot \frac{\left( \lambda t_0 \right)^m}{m!} \]
\[ = \left( \frac{2c}{t_0} \right)^m \]

**Conditional** \( m \mid N(t_0) = m \),
\[ F_{(m)}(x) = \left( \frac{2c}{t_0} \right)^m \]
\[ 0 \leq x \leq t_0 \]
\[ = 1 \quad , \quad x > t_0 \]
\[ = 0 \quad , \quad 0 \leq x \leq t_0 \]

\[ \therefore \]
\[
\hat{f}(x) = \frac{m}{t_0} \left( \frac{2x}{t_0} \right)^{m-1}, \quad 0 \leq x \leq t_0
\]
\[
= 0, \quad \text{otherwise}
\]

If \( t_0 = 1 \) then
\[
\hat{f}(x) = m \cdot x^{m-1}, \quad 0 \leq x \leq 1
\]
\[
= 0, \quad \text{otherwise}
\]

**Remark**

\[
E(Y|X=x) = \sum_y y \cdot f(y|x)
\]
\[
\sum_y f(y|x)
\]
discrete

**E(Y|X) = \mu(X)**

\[
E\left[ E(Y|X) \right] = E(Y)
\]

\[
E\left[ E(z|X) \right] = E(z)
\]

\[
\mu(x) = E(Y|X=x) = \int_0^\infty y \cdot f(y|x) \, dy
\]

\[
\text{cond pdf}
\]