Probability Theory

Textbook: A First Course in Probability, Sheldon Ross, 2019.

Prof. Hicham Elmongui

elmongui@alexu.edu.eg

Chapter 06: Jointly Distributed Random Variables

Joint Cumulative Probability Distribution Function

Definition

For any two random variables X and Y, the joint cumulative probability distribution function of X and Y is defined by

$$F(a,b) = P\{X \le a, Y \le b\} \qquad -\infty < a, b < \infty$$

Marginal CDF from joint CDF

$$\begin{aligned} F_X(a) &= P\{X \le a\} \\ &= P\{X \le a, Y < \infty\} \\ &= P\left(\lim_{b\to\infty} \{X \le a, Y \le b\}\right) \\ &= \lim_{b\to\infty} P\{X \le a, Y \le b\} \\ &= \lim_{b\to\infty} F(a,b) \\ &\equiv F(a,\infty) \end{aligned}$$

$$\begin{aligned} F_Y(b) &= P\{Y \le b\} \\ &= \lim_{a\to\infty} F(a,b) \\ &\equiv F(\infty,b) \end{aligned}$$

Joint Cumulative Probability Distribution Function

Answering probability statements in terms of joint CDF

$$\begin{aligned} P\{X > a, Y > b\} &= 1 - P\left(\{X > a, Y > b\}^c\right) \\ &= 1 - P\left(\{X > a\}^c \cup \{Y > b\}^c\right) \\ &= 1 - P\left(\{X \le a\} \cup \{Y \le b\}\right) \\ &= 1 - [P\{X \le a\} + P\{Y \le b\} - P\{X \le a, Y \le b\}] \\ &= 1 - F_X(a) - F_Y(b) + F(a,b) \end{aligned}$$

In general, whenever $a_1 < a_2$ and $b_1 < b_2$

$$P\{a_1 < X \le a_2, b_1 < Y \le b_2\} = F(a_2,b_2) + F(a_1,b_1) - F(a_1,b_2) - F(a_2,b_1)$$

Joint Probability Mass Function

Definition

When X and Y are both discrete RVs, the joint pmf of X and Y is defined by

$$p(x,y) = P\{X = x, Y = y\}$$

Marginal pmf from joint pmf

$$p_X(x) = P\{X = x\} = \sum_{y:p(x,y)>0} p(x,y)$$

$$p_Y(y) = P\{Y = y\} = \sum_{x:p(x,y)>0} p(x,y)$$

Example

Suppose that 3 balls are randomly selected from an urn containing 3 red, 4 white, and 5 blue balls. Let X and Y denote, respectively, the number of red and white balls chosen. Find the joint pmf of X and Y.

Solution

$$p(i,j) = \binom{3}{i}\binom{4}{j}\binom{5}{3-i-j}\Big/\binom{12}{3}$$

i \ j	0	1	2	3	Row sum $P\{X=i\}$
0	$10/220$	$40/220$	$30/220$	$4/220$	$84/220$
1	$30/220$	$60/220$	$18/220$	0	$108/220$
2	$15/220$	$12/220$	0	0	$27/220$
3	$1/220$	0	0	0	$1/220$
Col sum $P\{Y=j\}$	$56/220$	$112/220$	$48/220$	$4/220$	

Joint Probability Density Function

Definition

For any two *jointly continuous* RVs, X and Y, the joint pdf of X and Y is defined, over every set C in the two-dimensional plane, by

$$P\{(X,Y) \in C\} = \iint_{(x,y)\in C} f(x,y)\, dx\, dy$$

by defining $C = \{(x,y) : x \in A, y \in B\}$,

$$P\{X \in A, Y \in B\} = \int_B \int_A f(x,y)\, dx\, dy$$

$$F(a,b) = P\{X \le a, Y \le b\} = \int_{-\infty}^{b} \int_{-\infty}^{a} f(x,y)\, dx\, dy$$

Upon differentiation,

$$f(a,b) = \frac{\partial^2}{\partial a\, \partial b} F(a,b)$$

Examples

The joint density function of X and Y is given by

$$f(x,y) = c(y^2 - x^2)e^{-y} \qquad |x| \le y, \quad 0 < y < \infty$$

Find c

Solution

$$\begin{aligned}
1 &= \iint_{(x,y)} f(x,y)\ dx\,dy \\
&= \int_0^\infty \int_{-y}^{y} c(y^2 - x^2)e^{-y}\ dx\,dy \\
&= c\int_0^\infty e^{-y}\left(\left(xy^2 - \frac{x^3}{3}\right)\Big|_{-y}^{y}\right)\ dy \\
&= \frac{4c}{3}\int_0^\infty y^3 e^{-y}\ dy
\end{aligned}$$

$$\begin{aligned}
1 &= \frac{4c}{3} \times \Gamma(4) \\
&= \frac{4c}{3} \times 3! \\
&= 8c
\end{aligned}$$

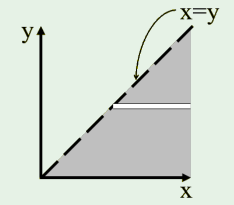
$$c = \frac{1}{8}$$

Examples (cont'd)

The joint pdf of X and Y is given by

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & 0 < x, y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Compute $P\{X > Y\}$



Solution

$$\begin{aligned}
P\{X > Y\} &= \iint_{(x,y):x>y} f(x,y)\ dx\,dy \\
&= \int_0^\infty \int_y^\infty 2e^{-x}e^{-2y}\ dx\,dy \\
&= \int_0^\infty 2e^{-2y}\left(-e^{-x}\Big|_y^\infty\right)\ dy \\
&= \int_0^\infty 2e^{-2y}e^{-y}\ dy \\
&= \int_0^\infty 2e^{-3y}\ dy \\
&= -\frac{2}{3}e^{-3y}\Big|_0^\infty \\
&= \frac{2}{3}
\end{aligned}$$

Marginal Probability Density Function

Definition

If X and Y are jointly continuous, they are individually continuous. Their pdf can be obtained as follows:

$$\begin{aligned}
P\{X \in A\} &= P\{X \in A, Y \in (-\infty, \infty)\} \\
&= \int_A \int_{-\infty}^{\infty} f(x,y)\ dy\,dx \\
&= \int_A f_X(x)\,dx
\end{aligned}$$

where

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\ dy$$

Marginal pdf from joint pdf

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\ dy \qquad f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\ dx$$

Example

The joint density function of X and Y is given by

$$f(x,y) = \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) \qquad 0 < x < 1, \quad 0 < y < 2$$

Compute the density function of X

Solution

$$\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f(x,y)\ dy \\
&= \int_0^2 \frac{6}{7}\left(x^2 + \frac{xy}{2}\right)\ dy \\
&= \frac{6}{7}\left(x^2 y + \frac{xy^2}{4}\right)\Big|_0^2 \\
&= \frac{6}{7}\left(2x^2 + x\right) \qquad 0 < x < 1
\end{aligned}$$

Independent Random Variables

Definition

The RVs X and Y are said to be *independent* if, for any two sets of real numbers A and B,

$$P\{X \in A,\ Y \in B\} = P\{X \in A\}P\{Y \in B\}$$

In other words, X and Y are independent if, for all A and B, the events $E_A = \{X \in A\}$ and $F_B = \{Y \in B\}$ are independent.

In terms of joint CDF, X and Y are independent if

$$F(a,b) = F_X(a)F_Y(b) \qquad \text{for all } a, b$$

In case of discrete RVs, X and Y are independent if

$$p(x,y) = p_X(x)p_Y(y) \qquad \text{for all } x, y$$

In case of continuous RVs, X and Y are independent if

$$f(x,y) = f_X(x)f_Y(y) \qquad \text{for all } x, y$$

Random variables that are not independent are said to be dependent.

Examples

If X_1 and X_2 are independent exponential RVs with respective parameters λ_1 and λ_2, find the distribution of $Z = X_1/X_2$.

Solution

Since X_1 and X_2 are independent, their joint pdf would be

$$f_{X_1,X_2}(x,y) = \lambda_1 e^{-\lambda_1 x}\lambda_2 e^{-\lambda_2 y}, \qquad x, y > 0$$

The distribution of $Z = X_1/X_2$ is

$$\begin{aligned}
F_Z(z) &= P(Z \le z) \\
&= P(X_1/X_2 \le z) \\
&= P(X_1 \le zX_2) \\
&= \iint_{(x,y):x \le zy} f_{X_1,X_2}(x,y)\ dx\,dy \\
&= \int_0^\infty \int_0^{zy} \lambda_1 e^{-\lambda_1 x}\lambda_2 e^{-\lambda_2 y}\ dx\,dy \quad = \frac{\lambda_1 z}{\lambda_1 z + \lambda_2}
\end{aligned}$$

Examples (cont'd)

The joint density function of X and Y is given by

$$f(x,y) = xe^{-(x+y)} \qquad x > 0, \quad y > 0$$

Are X and Y independent?

Solution

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x,y)\ \mathrm{d}y \\ &= \int_0^{\infty} xe^{-(x+y)}\ \mathrm{d}y \\ &= xe^{-x}\int_0^{\infty} e^{-y}\ \mathrm{d}y \\ &= xe^{-x}\left(-e^{-y}\Big|_0^{\infty}\right) \\ &= xe^{-x}, \quad x > 0 \end{aligned}$$

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)\ \mathrm{d}x \\ &= \int_0^{\infty} xe^{-(x+y)}\ \mathrm{d}x \\ &= e^{-y}\int_0^{\infty} xe^{-x}\ \mathrm{d}x \\ &= e^{-y}\int_0^{\infty} e^{-x}\ \mathrm{d}x \\ &= e^{-y}, \quad y > 0 \end{aligned}$$

$$f(x,y) = f_X(x)f_Y(y) \quad \forall x, y \qquad \Longrightarrow X \text{ and } Y \text{ are independent.}$$

Independent Random Variable

Proposition

The continuous (discrete) random variables X and Y are independent *if and only if* their joint pdf (pmf) can be expressed as

$$f_{X,Y}(x,y) = h(x)g(y) \qquad -\infty < x < \infty, \quad -\infty < y < \infty$$

Proof

Independence implies that the preceding factorization holds.

Now, suppose that $\quad f_{X,Y}(x,y) = h(x)g(y)$

then $\quad 1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y = \underbrace{\int_{-\infty}^{\infty} h(x)\,\mathrm{d}x}_{C1}\underbrace{\int_{-\infty}^{\infty} g(y)\,\mathrm{d}y}_{C_2} = C_1C_2$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y = C_2h(x), \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x = C_1g(y)$$

Since $C_1C_2 = 1$, it follows that $\quad f_{X,Y}(x,y) = f_X(x)f_Y(y)$ ■

Example

The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} 24xy & 0 < x < 1, \quad 0 < y < 1, \quad 0 < x+y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Are X and Y independent?

Solution

Because the region in which the joint density is nonzero cannot be expressed in the form $x \in A, y \in B$, the joint density does not factor, so the random variables are not independent.

Let

$$I(x,y) = \begin{cases} 1 & 0 < x < 1, \quad 0 < y < 1, \quad 0 < x+y < 1 \\ 0 & \text{otherwise} \end{cases}$$

The function $\quad f(x,y) = 24xy\ I(x,y) \quad$ clearly does not factor into a part depending only on x and another depending only on y, so the RVs are not independent.

Sums of Independent Random Variables

CDF of X+Y

$$\begin{aligned} F_{X+Y}(a) &= P\{X + Y \le a\} \\ &= \iint_{x+y\le a} f_X(x)f_Y(y)\ \mathrm{d}x\,\mathrm{d}y \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{a-y} f_X(x)f_Y(y)\ \mathrm{d}x\,\mathrm{d}y \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{a-y} f_X(x)\ \mathrm{d}x\ f_Y(y)\ \mathrm{d}y \\ &= \int_{-\infty}^{\infty} F_X(a-y)f_Y(y)\ \mathrm{d}y \end{aligned}$$

PDF of X+Y

$$\begin{aligned} f_{X+Y}(a) &= \tfrac{\mathrm{d}}{\mathrm{d}a}F_{X+Y}(a) \\ &= \tfrac{\mathrm{d}}{\mathrm{d}a}\int_{-\infty}^{\infty} F_X(a-y)f_Y(y)\ \mathrm{d}y \\ &= \int_{-\infty}^{\infty} \tfrac{\mathrm{d}}{\mathrm{d}a}F_X(a-y)f_Y(y)\ \mathrm{d}y \\ &= \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)\ \mathrm{d}y \end{aligned}$$

The pdf f_{X+Y} is the convolution of the pdf's of X and Y.

Sums of i.i.d. Uniform Random Variables

If X and Y are independent random variables, both uniformly distributed on (0, 1), calculate the probability density of $X + Y$.

Solution

$$f_X(a) = f_Y(a) = \begin{cases} 1 & 0 < a < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} f_{X+Y}(a) &= \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)\ \mathrm{d}y \\ &= \int_0^1 f_X(a-y)\ \mathrm{d}y \end{aligned}$$

For $0 < a \le 1$,

$$f_{X+Y}(a) = \int_0^a 1\ \mathrm{d}y + \int_a^1 0\ \mathrm{d}y = a$$

For $1 < a < 2$,

$$f_{X+Y}(a) = \int_0^{a-1} 0\ \mathrm{d}y + \int_{a-1}^1 1\ \mathrm{d}y = 2 - a$$

$$f_{X+Y}(a) = \begin{cases} a & 0 < a \le 1 \\ 2-a & 1 < a < 2 \end{cases}$$

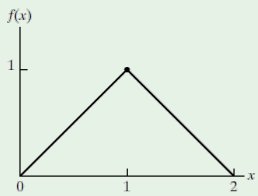


Figure: Triangular Distribution

Sums of Independent Gamma Random Variables

If X and Y are independent gamma random variables with respective parameters (s, λ) and (t, λ), calculate the probability density of $X + Y$.

Solution

$$f_X(x) = \frac{\lambda e^{-\lambda x}(\lambda x)^{s-1}}{\Gamma(s)} \qquad f_Y(y) = \frac{\lambda e^{-\lambda y}(\lambda y)^{t-1}}{\Gamma(t)}$$

$$\begin{aligned} f_{X+Y}(a) &= \int_0^a \frac{\lambda e^{-\lambda(a-y)}[\lambda(a-y)]^{s-1}}{\Gamma(s)} \times \frac{\lambda e^{-\lambda y}(\lambda y)^{t-1}}{\Gamma(t)}\ \mathrm{d}y \\ &= \frac{\lambda^{s+t}\, e^{-\lambda a}}{\Gamma(s)\Gamma(t)}\int_0^a (a-y)^{s-1}y^{t-1}\ \mathrm{d}y \\ &= \frac{\lambda^{s+t}\, e^{-\lambda a}\ a^{s+t-1}}{\Gamma(s)\Gamma(t)}\int_0^1 (1-z)^{s-1}z^{t-1}\ \mathrm{d}z \qquad \text{(by letting } z = y/a\text{)} \\ &= \frac{\lambda e^{-\lambda a}(\lambda a)^{s+t-1}}{\Gamma(s)\Gamma(t)} \times B(s,t) \qquad = \frac{\lambda e^{-\lambda a}(\lambda a)^{s+t-1}}{\Gamma(s)\Gamma(t)} \times \frac{\Gamma(s)\Gamma(t)}{\Gamma(s+t)} \\ &= \frac{\lambda e^{-\lambda a}(\lambda a)^{s+t-1}}{\Gamma(s+t)} \qquad \Longrightarrow X + Y \sim \text{Gamma}(s+t, \lambda) \end{aligned}$$

Sums of Independent Gamma Random Variables

For a fixed value of λ, the family of gamma distributions is closed under convolution.

Proposition

If $X_i, i = 1, 2, \cdots, n$ are independent gamma random variables with respective parameters $(t_i, \lambda), i = 1, 2, \cdots, n$, then

$$\sum_{i=1}^{n} X_i \sim \text{Gamma}\left(\sum_{i=1}^{n} t_i, \lambda\right)$$

Example

If $X_i, i = 1, 2, \cdots, n$ are i.i.d. exponential RVs with parameter λ, calculate the pdf of $X_1 + X_2 + \cdots + X_n$.

Solution

An exponential RV with parameter λ is the same as a gamma RV with parameters $(1, \lambda)$, it follows that $X_1 + X_2 + \cdots + X_n \sim \text{Gamma}(n, \lambda)$.

Example: χ^2 Probability Distribution

If $Z_1, Z_2, \cdots, Z_n$ are indep standard normal RVs, then $\chi_n^2 = \sum_{i=1}^n Z_i^2$ is said to have the *chi-squared* dist. with n degrees of freedom. Calculate $f_{\chi_n^2}(y)$

Solution

$$f_{Z^2}(y) = \frac{1}{2\sqrt{y}}\left(f_Z(\sqrt{y}) + f_Z(-\sqrt{y})\right) = \frac{1}{2\sqrt{y}} \times \frac{2}{\sqrt{2\pi}} e^{-y/2} = \frac{(1/2)^{\frac{1}{2}} y^{\frac{1}{2}-1} e^{-y/2}}{\Gamma\left(\frac{1}{2}\right)}$$

$$\Longrightarrow \chi_1^2 = Z^2 \sim \text{Gamma}\left(\tfrac{1}{2}, \tfrac{1}{2}\right)$$

$$\chi_n^2 = \sum_{i=1}^{n} Z_i^2$$

$$\chi_n^2 \sim \text{Gamma}\left(\tfrac{n}{2}, \tfrac{1}{2}\right)$$

$$f_{\chi_n^2}(y) = \frac{(1/2)^{\frac{n}{2}} y^{\frac{n}{2}-1} e^{-y/2}}{\Gamma\left(\frac{n}{2}\right)}, \quad y > 0$$

$$= \frac{y^{\frac{n}{2}-1} e^{-y/2}}{2^{\frac{n}{2}}\, \Gamma\left(\frac{n}{2}\right)}, \quad y > 0$$

$$\Gamma\left(\frac{n}{2}\right) = \begin{cases} \left(\frac{n}{2}-1\right)! & n \text{ is even} \\ \left(\frac{n}{2}-1\right)\Gamma\left(\frac{n}{2}-1\right) & n \text{ is odd} \end{cases}$$

Sums of Independent Normal Random Variables

If X and Y are independent RVs that are normally distributed with respective parameters (μ_1, σ_1^2) and (μ_2, σ_2^2), calculate the pdf of $X + Y$.

Solution

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right\}, \quad f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left\{-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right\}$$

$$f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y) f_Y(y)\, dy$$

$$f_{X+Y}(a) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{-\frac{((a-y)-\mu_1)^2}{2\sigma_1^2}\right\} \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left\{-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right\} dy$$

$$= \cdots = \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}} \exp\left\{-\frac{(a-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}\right\}$$

$$X + Y \sim \text{Normal}\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\right)$$

Sums of Independent Normal Random Variables

The family of normal distributions is closed under convolution.

Proposition

If $X_i, i = 1, 2, \cdots, n$ are independent normal random variables with respective parameters $(\mu_i, \sigma_i^2), i = 1, 2, \cdots, n$, then

$$\sum_{i=1}^{n} X_i \sim \text{Normal}\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$$

Example

A basketball team will play a 44-game season of which 26(18) games against class A(B) teams. Suppose that the team will beat a class A(B) team with probability .4(.7) and that the results of the different games are independent. How likely is it that the team wins 25 games or more?

Solution

- Let X_A and X_B respectively denote the number of games the team wins against class A and against class B teams.
- X_A and X_B are indep binomial RVs (approximated as normal RVs)

$$E[X_A] = 26(.4) = 10.4$$
$$\text{Var}(X_A) = 26(.4)(.6) = 6.24$$
$$E[X_B] = 18(.7) = 12.6$$
$$\text{Var}(X_B) = 18(.7)(.3) = 3.78$$
$$E[X_A + X_B] = 10.4 + 12.6 = 23$$
$$\text{Var}(X_A + X_B) = 6.24 + 3.78 = 10.02$$

$P\{X_A + X_B \geq 25\}$

$= P\{X_A + X_B \geq 24.5\}$ (cont. correction)

$$= P\left\{\frac{X_A + X_B - 23}{\sqrt{10.02}} \geq \frac{24.5 - 23}{\sqrt{10.02}}\right\}$$

$\approx P\{Z \geq .4739\} = 1 - P\{Z < .4739\}$

$\approx .3178$

Examples (cont'd)

Starting at some fixed time, let $S(n)$ denote the price of a certain security at the end of n additional weeks, $n \geq 1$. A popular model for the evolution of these prices assumes that the price ratios $S(n)/S(n-1), n \geq 1$, are i.i.d. lognormal random variables. Assuming this model, with parameters $\mu = .0165, \sigma = .0730$, what is the probability that the price of the security increases over each of the next two weeks?

Y is *lognormal* if it can be expressed as $Y = e^X$ where X is a normal RV.

Solution

The probability that the price is up after one week is

$$P\left\{\frac{S(1)}{S(0)} > 1\right\} = P\left\{\log\left(\frac{S(1)}{S(0)}\right) > 0\right\} = P\left\{Z > \frac{0 - .0165}{.0730}\right\}$$

$$\approx P\{Z > -.2260\} = P\{Z < .2260\} \approx .5894$$

Since the successive price ratios are independent, the probability that the price increases over each of the next two weeks is $(.5894)^2 = .3474$

Sums of Independent Poisson Random Variables

If X and Y are independent Poisson random variables with respective parameters λ_1 and λ_2, calculate the distribution of $X+Y$.

Solution

The event $\{X+Y=n\}$ may be written as the union of the disjoint events $\{X=k,Y=n-k\}, 0\le k\le n$. Therefore,

$$P\{X+Y=n\} = \sum_{k=0}^{n} P\{X=k,Y=n-k\} = \sum_{k=0}^{n} P\{X=k\}P\{Y=n-k\} = \sum_{k=0}^{n} \mathrm{e}^{-\lambda_1}\frac{\lambda_1^k}{k!}\times \mathrm{e}^{-\lambda_2}\frac{\lambda_2^{n-k}}{(n-k)!}$$

$$= \mathrm{e}^{-(\lambda_1+\lambda_2)}\sum_{k=0}^{n}\frac{1}{k!(n-k)!}\lambda_1^k\lambda_2^{n-k} = \frac{\mathrm{e}^{-(\lambda_1+\lambda_2)}}{n!}\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}\lambda_1^k\lambda_2^{n-k} = \mathrm{e}^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^n}{n!}$$

$$X+Y\sim \mathrm{Poisson}(\lambda_1+\lambda_2)$$

If $X_i, i=1,2,\cdots,n$ are independent Poisson RVs with respective parameters $\lambda_i, i=1,2,\cdots,n$, then $X_1+X_2+\cdots+X_n\sim \mathrm{Poisson}(\lambda_1+\lambda_2+\cdots+\lambda_n)$

Sums of Independent Binomial Random Variables

If X and Y are independent binomial random variables with respective parameters (n,p) and (m,p), calculate the distribution of $X+Y$.

Solution

Recalling the interpretation of a binomial RV, and without any computation at all, $X+Y\sim \mathrm{Binomial}(n+m,p)$ ■

$$P\{X+Y=k\} = \sum_{i=0}^{n} P\{X=i,Y=k-i\} = \sum_{i=0}^{n} P\{X=i\}P\{Y=k-i\} = \sum_{i=0}^{n}\binom{n}{i}p^iq^{n-i}\times\binom{m}{k-i}p^{k-i}q^{m-k+i}$$

$$= p^kq^{m+n-k}\sum_{i=0}^{n}\binom{n}{i}\binom{m}{k-i} = \binom{n+m}{k}p^kq^{m+n-k}$$

$$X+Y\sim \mathrm{Binomial}(n+m,p)$$

If $X_i, i=1,\cdots,N$ are independent binomial RVs with respective parameters $(n_i,p), i=1,\cdots,N$, then $X_1+\cdots+X_N\sim \mathrm{Binomial}(n_1+\cdots+n_N,p)$