Recall :-

- Conditional Probability
- Independence

$X_1, X_2, \ldots, X_n$ are i.i.d $\equiv \begin{cases} i - \text{independent} \\ i - \text{identically} \\ d - \text{distributed} \end{cases}$

[Discrete r.v.(s)]

$$\mathbb{P}(X_1 \leq x) = \mathbb{P}(X_i \leq x) \quad \forall i \geq 2$$

and $\mathbb{P}(X_{i_1} = x_1, X_{i_2} = x_2, \ldots X_{i_n} = x_n)$

$$\{x_i \in \mathbb{R} : i \geq 1\} \qquad = \prod_{j=1}^{n} \mathbb{P}(X_{i_j} = x_j)$$

[Discrete r.v.(s)] - Joint distribution $(X_1, \ldots, X_n)$

$$\{\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n)\}_{x_i \in \text{Range}(X_i)}$$

Sums of Random Variables

Recall:

Let $X, Y \sim \text{Bernoulli}(p)$ be two independent random variables.

$Z = X + Y$, $Z \sim ?$. $Z \sim \text{Binomial}(2, p)$

Inductively:-

- $X_1, X_2, \ldots, X_n$ are i.i.d Bernoulli (p)

$$\Rightarrow \quad X = \sum_{i=1}^{n} X_i \sim \text{Binomial}(n, p)$$

Sums of Random Variables

Let $X \sim \text{Poisson}(\lambda_1)$ and $Y \sim \text{Poisson}(\lambda_2)$ be independent random variables.

a. Let $Z = X + Y$. Find the distribution of Z.

b. Find the conditional distribution of $X \mid Z$.

---

a. $X \sim \text{Poisson}(\lambda_1) \Rightarrow \text{Range}(X) \in \{0,1,2,\dots\}$

$Y \sim \text{Poisson}(\lambda_2) \Rightarrow \text{Range}(Y) \in \{0,1,2,\dots\}$

$\Rightarrow Z = X + Y$ has $\text{Range}(Z) = \{0,1,2,\dots\}$

Take $z \in \text{Range}(Z)$

$$\mathbb{P}(Z=z) = \mathbb{P}(X+Y=z) = \mathbb{P}\left(\bigcup_{i=0}^{z} (X=i,\ Y=z-i)\right)$$

$$= \sum_{i=0}^{z} \mathbb{P}(X=i,\ Y=z-i)$$

Mutually Exclusive events

$$= \sum_{i=0}^{z} \mathbb{P}(X=i)\ \mathbb{P}(Y=z-i)$$

X and Y are independent

$$= \sum_{i=0}^{z} \frac{e^{-\lambda_1} \lambda_1^{i}}{i!} \frac{e^{-\lambda_2} (\lambda_2)^{z-i}}{(z-i)!}$$

$X \sim \text{Poisson}(\lambda_1)$

$Y \sim \text{Poisson}(\lambda_2)$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{z!} \sum_{i=0}^{z} \frac{\lambda_1^{i} (\lambda_2)^{z-i}}{i!\,(z-i)!} \cdot z!$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{z!} (\lambda_1+\lambda_2)^{z}$$

If $X \sim \text{Poisson}(\lambda_1)$ and $Y \sim \text{Poisson}(\lambda_2)$ then

$$Z \sim \text{Poisson}(\lambda_1+\lambda_2)$$

Inductively : $X_i \sim \text{Poisson}(\lambda_i)$ $i = 1, 2, \ldots, n$ and independent

$Z = \sum_{i=1} X_i$ then $Z \sim \text{Poisson}\left(\sum_{i=1}^{n} \lambda_i\right)$

(b) Find distribution of $X \mid Z$

Fix $z \in \text{Range}(Z)$

Step 1 : $X \mid Z = z$ — Find its distribution

$x \in \text{Range}(X)$

$$\mathbb{P}(X = x \mid Z = z) = ? \quad \begin{cases} 0 & x > z \\ ? & x \leq z \end{cases}$$

For $0 \leq x \leq z \Rightarrow$

$$= \frac{\mathbb{P}(X = x, Z = z)}{\mathbb{P}(Z = z)}$$

$$= \frac{\mathbb{P}(X = x, Y = z - x)}{\mathbb{P}(Z = z)}$$

independence of X and Y

$$= \frac{\mathbb{P}(X = x)\, \mathbb{P}(Y = z - x)}{\mathbb{P}(Z = z)}$$

$$= \frac{\dfrac{e^{-\lambda_1} \lambda_1^{x}}{x!} \; \dfrac{e^{-\lambda_2} \lambda_2^{z-x}}{z-x!}}{\dfrac{e^{-(\lambda_1+\lambda_2)} (\lambda_1+\lambda_2)^{z}}{z!}}$$

$$= \frac{z!}{x!\,(z-x)!} \; \frac{\lambda_1^{x} \; \lambda_2^{z-x}}{(\lambda_1+\lambda_2)^{z}}$$

$$\mathbb{P}(X=x \mid Z=z) = {}^{z}C_{x}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{x}\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{z-x}$$

$x = 0, 1, 2, \ldots, z$

independent $\qquad$ $Z \sim \text{Poisson}(\lambda_1+\lambda_2)$

$X \sim \text{Poisson}(\lambda_1) \qquad Y \sim \text{Poisson}(\lambda_2) \qquad Z = X+Y$

$$\Rightarrow \quad X \mid Z=z \sim \text{Binomial}\left(z, \frac{\lambda_1}{\lambda_1+\lambda_2}\right)$$

$$\boxed{X \mid Z \sim \text{Binomial}\left(Z, \frac{\lambda_1}{\lambda_1+\lambda_2}\right)}$$

Expectation :- X, Y are random variables

$$E[X+Y] = E[X] + E[Y]$$

Q: $\quad E[XY] = ?$ $\quad E[X]$ and $E[Y]$

Expectation of product

Ex:- Construct X, Y random variables

$E[XY] \neq E[X]E[Y]$.

In general $E[XY] \neq E[X]E[Y]$ but

If X and Y are independent, then $E[XY] = E[X]\,E[Y]$

$$E[XY] = \sum_{t \in \text{Range}(XY)} t \cdot \mathbb{P}(XY = t)$$

$$= \sum_{\substack{u \in \text{Range}(X) \\ v \in \text{Range}(Y)}} uv\, \mathbb{P}(X = u, Y = v)$$

independence of X and Y $= \sum_{\substack{u \in \text{Range}(X) \\ v \in \text{Range}(Y)}} u\, v\ \mathbb{P}(X=u)\, \mathbb{P}(Y=v)$

$$= \sum_{u \in \text{Range}(X)} u\ \mathbb{P}(X=u) \sum_{v \in \text{Range}(Y)} v\ \mathbb{P}(Y=v)$$

$$= E[X]\ E[Y]$$

D

X, Y Independent $\Rightarrow$ $E[XY] - E[X]\,E[Y] = 0$

$$\boxed{\text{"Dependence" measure} \quad E[(X - E[X])(Y - E[Y])] = ?}$$

$$E(X - E[X])(Y - E[Y]) = E\,XY - E[X]\,E[Y] - E[Y]\,E[X] + E[X]\,E[Y]$$

$$= E[XY] - E[X]\,E[Y]$$

Ex: Construct random variables

(i) $E[XY] - E[X]\,E[Y] = 0$

(ii) X and Y are NOT independent

Covariance - measure of dependence between X and Y

If X and Y are independent $\Rightarrow$ $Cov[X,Y] = 0$

The "covariance of X and Y" is defined as
$Cov[X, Y] = E[(X - E[X])(Y - E[Y])]$.

$Cov[X,Y] > 0 \equiv \begin{cases} X - \text{larger than its average} \mid \text{smaller than its average} \\ \quad \text{at the same time} \\ Y - \text{larger than its average} \mid \text{smaller than its average} \end{cases}$

Similar interpretation one can do for $Cov[X,Y] < 0$.

Covariance-Facts

$Cov[X, Y] = Cov[Y, X]$; – Symmetric

$Cov[X, aY + bZ] = a \cdot Cov[X, Y] + b \cdot Cov[X, Z]$; – linear

$Cov[aX + bY, Z] = a \cdot Cov[X, Z] + b \cdot Cov[Y, Z]$; and

If X and Y are independent with a finite covariance, then $Cov[X, Y] = 0$.

$Var[X + Y] = Var[X] + Var[Y] + 2Cov[X, Y]$.

$= E[((X+Y) - E[X+Y])^2]$

$\equiv$ expand & do algebra

Conditional Variance

Theorem 4.4.9

Let X and $Y : S \to T$ be two discrete random variables on a sample space S. let $g(y) = E[X|Y = y]$. Let $h(y) = Var[X|Y = y]$. Denoting $g(Y)$ by $E[X|Y]$ and denoting $h(Y)$ by $Var[X|Y]$, then

$$Var[X] = E[Var[X|Y]] + Var[E[X|Y]].$$

and: $E[X] = E[E[X|Y]]$

Conditional Variance

Question: The number of eggs N found in nests of a certain species of turtles has a Poisson distribution with mean λ. Each egg has probability p of being viable and this event is independent from egg to egg. Find the mean and variance of the number of viable eggs per nest.

N = number of eggs

X = number of viable ones.

$N = n$ for some $n \geq 1$

$$X \mid N = n \sim \text{Binomial}(n, p)$$

$$\Rightarrow \quad g(n) := E[X \mid N = n] = np$$

$$h(n) := Var[X \mid N = n] = np(1-p)$$

$$\Rightarrow \quad \underset{\underset{E[X|N]}{|||}}{g(N)} = Np \quad \& \quad \underset{\underset{Var[X|N]}{|||}}{h(N)} = Np(1-p)$$

- $E[X] = E[E[X \mid N]] = E[pN] = p\,E[N]$
$$= p\lambda$$

- Analysis of Variance $\quad Var[X] = E[Var[X|Y]] + Var[E[X|Y]].$

- $Var[X] = E[h(N)] + Var[g(N)]$
$$= E[N p(1-p)] + Var[Np]$$
$$= p(1-p)\,E[N] + p^2\,Var[N]$$

$N \sim$ Poisson(λ)
$$= p(1-p)\lambda + p^2\lambda = p\lambda$$