Conditional Problablity

Let the problality set function $P(C)$ be defined on the sample space $\mathcal{C}$ and C_1 be a subset of $\mathcal{C}$ such that $P(C_1) > 0$.

we agree to consider only thoes outcome of Random experiment that are elements of C_1 (basically we are going to take elements of C_1 to be the sample space.)

This probality is denoted by $P(C_2|C_1)$ where $C_2 \subseteq \mathcal{C}$

$$P(C_2|C_1) = \frac{P(C_1 \cap C_2)}{P(C_1)}$$

$$P(C_1|C_1) = 1$$

$$P(C_2 \cap C_1|C_1) = P(C_2|C_1)$$

$$\cancel{\frac{P(C_2 \cap C_1 | C_1)}{P(C_1 \cap C_1)} = \frac{P(C_2 \cap C_1)}{P(C_1)}}$$

Properties :-

$$P(C_2|C_1) \geq 0$$

$$P(C_2 \cup C_3 \cdots | C_1) = P(C_2|C_1) + P(C_3|C_1) \cdots$$

if $C_2, C_3 \cdots$ are disjoint

$$P(C_1|C_1) = 1$$

Marginal Distribution

Suppose $f(x_1, x_2)$ be pdf of 2 RV X_1 & X_2

Consider 2 event $a < X_1 < b$ & $-\infty < X_2 < \infty$

and $a < X_1 < b$ these two events are identical

$\therefore$

$$Pr(a < X_1 < b) = Pr(a < X_1 < b, -\infty < X_2 < \infty)$$
$$= \int_a^b \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2\, dx_1$$

$\therefore$

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2$$

marginal probability of pdf X_1

Now define Let X_1 & X_2 be discrete RV $A_1 = \{(x_1, x_2) : x_1 = x_1'\}$ $A_2 = \{(x_1, x_2) : x_2 = x_2'\}$

$$\therefore P(A_2 | A_1) = \frac{Pr(A_1 \cap A_2)}{Pr(A_1)} = \frac{Pr(X_1 = x_1', X_2 = x_2')}{Pr(X_1 = x_1')} = \frac{f(x_1, x_2)}{f(x_1)}$$

Same can be done for cts case

$$f(x_2 | x_1) = \frac{f(x_1, x_2)}{f(x_1)}$$

The coordelation cofficient

Let X, Y be two RV with joint distribution $f(x, y)$

$\therefore$

$$\sigma_{12} = E[(x-u_1)(y-u_2)]$$

where u_1 = mean of RV X = $E[X]$

u_2 = mean of RV Y = $E[Y]$

From Linearity of E =

$$\sigma_{12} = E[xy] - u_1 u_2$$

~~correlation~~ covariance of X & Y

$$\text{Correlation} = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}$$

$-1 < \rho < 1$

$\Rightarrow$ Let X and Y be two RV

$\mathrm{Var}(aX + Y) \geq 0$

$$\Rightarrow a^2 \mathrm{Var}(X) + 2a\,\mathrm{Cov}(X,Y) + \mathrm{Var}(Y) \geq 0 \quad \forall a$$

$\therefore$ discriminant ≤ 0

$$\Rightarrow 4\,\mathrm{Cov}(X,Y)^2 - 4\,\mathrm{Var}(X)\,\mathrm{Var}(Y) \leq 0$$

$$\Rightarrow -1 \leq \rho^2 \leq 1$$

Change of Variable

Some time we ~~ar~~ are given a ~~pdf~~ RV X and we need to know the distribution of $f(x)$.

Notice that as X is an RV $f(x)$ will also be an RV.

There are various methods to find pdf or pmf of $f(x)$.

For Continous distributions

1) Using CDF :- We know that CDF of a distribution is unique and we can find the ~~pdf~~ pdf just by differenciating the CDF.

How to approach with ~~pdf~~ CDF.

Suppose X be any CRV with pdf $f(x)$

~~Let~~ You have to find distribution of $g(x)$

$$\therefore \Pr(g(x) \leq t) = F_{g(x)}(t)$$

This represents the CDF of $g(x)$

Now we can ~~expr~~ solve $\Pr(g(x) \leq t)$ in term of X and get the answer.

Lets take an example.

$X \sim N(0,1)$

we have to find distibution of x^2

$$\therefore \quad F_{x^2}(t) = \Pr(x^2 \leq t)$$

$$= \Pr(-\sqrt{t} \leq X \leq \sqrt{t})$$

$$F_{x^2}(t) = \int_{-\sqrt{t}}^{\sqrt{t}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

Now differenciate to get the pdf.

$$\frac{d}{dt} F_{x^2}(t) = \frac{1}{\sqrt{2\pi}} e^{-t/2} \frac{1}{2\sqrt{t}} + \frac{1}{2\sqrt{t}} \frac{e^{-t/2}}{\sqrt{2\pi}} = \frac{e^{-t/2}}{\sqrt{2\pi t}}$$

Interesting fact this is chi square distributio with degree of fredom 1.

This above method is universal and work in every case. Remember it's not the easiest to apply But when applied correcty always give the correct result.

Example suppose X & Y are both iid exponential RV with parameters & λ_2

Find distribution of $X+Y$

$$F_{x+y}(t) = \Pr(x+y \leq t)$$

$$f(x,y) = \lambda_2 \lambda_1 e^{-\lambda_1 x - \lambda_2 y}$$

$$Pr(X+Y \leq t) = \int_0^t \int_0^{t-y} \lambda_2 \lambda_1 e^{-\lambda_1 x} \cdot e^{-\lambda_2 y} \, dx \, dy$$

$$\int_0^{t-y} e^{-\lambda_1 x} dx = \frac{e^{-\lambda_1 x}}{-\lambda_1} \Big|_{t-y}^{0} = \frac{1 - e^{-\lambda_1 (t-y)}}{\lambda_1}$$

$$\lambda_2 \int_0^t \frac{1 - e^{-\lambda_1 (t-y)}}{} e^{-\lambda_2 y} dy$$

$$\lambda_2 \int_0^t e^{-\lambda_2 y} - e^{-\lambda_1 t + \lambda_1 y - \lambda_2 y} dy$$

$$\lambda_2 \left[\frac{e^{-\lambda_2 y}}{-\lambda_2} \Big|_t^0 + \frac{e^{-\lambda_1 t + (\lambda_1 - \lambda_2) y}}{\lambda_1 - \lambda_2} \Big|_t^0 \right]$$

$$1 - e^{-\lambda_2 t} + \frac{\lambda_2}{\lambda_1 - \lambda_2} e^{-\lambda_1 t} - e^{\cancel{-\lambda_1 t + \lambda_1 t} - \lambda_2 t}$$

$$F_{x+y}(t) = 1 - e^{-\lambda_2 t} + \frac{\lambda_2}{\lambda_1 - \lambda_2} \left[e^{-\lambda_1 t} - e^{-\lambda_2 t} \right]$$

$$\frac{d F_{x+y}(t)}{dt} = \lambda_2 e^{-\lambda_2 t} + \frac{\lambda_2}{\lambda_1 - \lambda_2} \left[-\lambda_1 e^{-\lambda_1 t} + \lambda_2 e^{-\lambda_2 t} \right]$$

But if $\lambda_1 = \lambda_2$ then

$$F_{x+y}(t) = 1 - e^{-\lambda t} - \lambda t e^{-\lambda t}$$

$$f_{x+y}(t) = \lambda e^{-\lambda t} - \lambda \left[e^{-\lambda t} - \lambda t e^{-\lambda t} \right] = \lambda^2 t e^{-\lambda t}$$

This distribution is Erlang $(2, \lambda)$

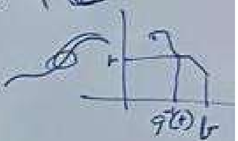
2) Find pdf with help of change of formula technique.

Let X be a RV.

Suppose $g(x)$ be a function st $g'(t)$ is non zero in an interval (a,b).

$\therefore$ $g(x)$ can be increasing or decreasing.

Then $$F_{g(x)}(t) = \Pr(g(x) \le t) = \sum_{\substack{\text{all such } (a,b) \text{ st } g'(t) \text{ is non} \\ \text{zero in } (a,b)}} \Pr(X \in (a,b) \mid g(x) \le t \text{ and } t \in (a,b)) + C$$



Increasing case

$$\sum_{(a,b)} \Pr(a < X \le g^{-1}(t))$$

$$\sum_{(a,b)} F_X(g^{-1}(t)) - F_X(a)$$

Decreasing case

$$\sum_{(a,b)} \Pr(b \ge X \ge g^{-1}(t))$$

$$\sum_{(a,b)} F_X(b) - F_X(g^{-1}(t))$$

$$f_{g(x)}(t) = \sum_{(a,b)} f_X(g^{-1}(t))\,(g^{-1}(t))' \quad \text{or} \quad \sum_{(a,b)} -f_X(g^{-1}(t))\,(g^{-1}(t))'$$

$$= \sum_{(a,b)} f_X(g^{-1}(t))\,|(g^{-1}(t))'|$$

Suppose $X \sim N(0,1)$ and $Y = X^2$ find pdf of Y

$\therefore$ ~~$f_Y(t) = \ldots_{(-\infty,\infty)}$~~

$g(x) = x^2$

$\therefore \quad g^{-1}(t) = -\sqrt{t} \quad$ if $x \in (-\infty, 0) \qquad g^{-1}(t)' = \frac{1}{2}\frac{-1}{\sqrt{t}}$

$g^{-1}(t) = \sqrt{t} \quad$ if $x \in (0, \infty) \qquad g^{-1}(t) = \frac{1}{2\sqrt{t}}$

$$\therefore \quad f_Y(t) = f_X(-\sqrt{t})\frac{1}{2\sqrt{t}} + f_X(\sqrt{t})\frac{1}{2\sqrt{t}}$$

$$= \frac{1}{\sqrt{t}} f_X(\sqrt{t}) = \frac{1}{\sqrt{2\pi t}} e^{-t/2}$$

In Case of ~~Two~~ ~~biva~~ multivariet distributions :-

$$T\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix}$$

$$\cancel{f_{A,B}(a,b)} = f_{A,B}(a,b) = f_{X,Y}(T^{-1}(a,b)) \, || J_{T^{-1}}(a,b) ||$$

Suppose X and Y ~~be~~ such that $f(x,y) = 1$ $0 \le x \le 1$ ~~and~~ $y \le 1$

Find $X+Y$, $X-Y$

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{matrix} X+Y \\ X-Y \end{matrix} = \begin{matrix} A \\ B \end{matrix} \qquad f_{A,B}(a,b) = f_{X,Y}\left(\frac{A+B}{2}, \frac{A-B}{2}\right)\frac{1}{2}$$

$$T^{-1}\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} \frac{A+B}{2} \\ \frac{A-B}{2} \end{pmatrix} \qquad f_{(A,B)}(a,b) = \frac{1}{2} \quad \begin{matrix} 0 \le \frac{A+B}{2} \le 1 \\ 0 \le \frac{A-B}{2} \le 1 \end{matrix}$$

$$J_{T^{-1}} \left|\left| \begin{matrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{matrix} \right|\right| = \frac{1}{2}$$