Probability & Statistics.

L-13

$X \to$ continuous RV

pdf : $f(x) \geq 0$, $\int_{-\infty}^{\infty} f(x)\,dx = 1$

PDF : Probability distribution fn.

$F_X(t) = P(X \leq t)$, $t \in \mathbb{R}$

$E(X)$, $V(X)$, moments "μ_n"

Moment generating function (mgf)

$$g(t) = E(e^{tX}) < \infty.$$

$$= \int_{-\infty}^{\infty} e^{tx} f_X(x)\,dx$$

$< \infty$.

$$\mu_n = E(X^n) = \int_{-\infty}^{\infty} x^n f_X(x)\,dx$$

$$\boxed{\mu_n = g^{(n)}(0)}$$

Ex. Let $X \sim U[0,1]$

$$g(t) = \int_0^1 e^{tx}\, dx$$

$$= \frac{e^t - 1}{t}$$

Ex. $X \sim E_\lambda$, $p_X(x) = \lambda e^{-\lambda x}$, $0 < x < \infty$

$$g(t) = \int_0^\infty e^{tx} \cdot \lambda e^{-\lambda x}\, dx$$

$$= \lambda \int_0^\infty e^{(t-\lambda)x}\, dx$$

$$= \frac{\lambda}{\lambda - t}$$

Ex. $X \sim N(0,1)$

$$g(t) = \frac{1}{\sqrt{2\pi}} \int e^{tx} \cdot e^{-x^2/2}\, dx$$

$$= e^{t^2/2}$$

[Propo]sition: If X is a bdd continuous [ra]ndom variable then the mgf $g(t)$ [d]etermines the density f_X uniquely.

Functions of continuous RV.

$$Y = aX + b$$

Q. What about the distribution fn. of Y i.e. $F_Y(t)$ given $F_X(t)$!!

$F_X(t) = P(X \le t)$

Case I $\quad a > 0$

$$\begin{aligned} F_Y(t) &= P(Y \le t) \\ &= P(aX + b \le t) \\ &= P\left(X \le \frac{t-b}{a}\right) \\ &= F_X\left(\frac{t-b}{a}\right) \end{aligned}$$

Case II $\quad a < 0$

$$\begin{aligned} F_Y(t) &= P(Y \le t) \\ &= P\left(X \ge \frac{t-b}{a}\right) \\ &= 1 - P\left(X < \frac{t-b}{a}\right) \end{aligned}$$

$P\left(X = \frac{t-b}{a}\right) = 0$

Assuming the continuity.

$$\begin{aligned} F_Y(t) &= 1 - P\left(X \le \frac{t-b}{a}\right) \\ &= 1 - F_X\left(\frac{t-b}{a}\right) \end{aligned}$$

Q. Given $p_X(x)$, what is

$p_Y(y)$ when $Y = aX + b$

Case I $\quad a > 0$

$$F_Y(t) = \cancel{P(Y \le t)}$$

$$= F_X\left(\frac{t-b}{a}\right)$$

$$= \int_{-\infty}^{\frac{t-b}{a}} p_X(x)\, dx$$

$$= \cancel{\int^{t} p\, q(y)\, dy}$$

$$= \int_{-\infty}^{t} \frac{1}{a}\, p\left(\frac{y-b}{a}\right) dy$$

$$= \int_{-\infty}^{t} q(y)\, dy$$

Then $q(y) = \frac{1}{a}\, p_X\left(\frac{t-b}{a}\right), \; t \in \mathbb{R}$.

Case II $\quad a < 0$. $\quad F_Y(t) = 1 - F_X\left(\frac{t-b}{a}\right)$

$$= \int_{\frac{t-b}{a}}^{\infty} p_X(x)\, dx$$

$$= \int_{-\infty}^{t} \frac{1}{|a|}\, p\left(\frac{y-b}{a}\right) dy = \int_{-\infty}^{t} q(y)\, dy$$

Theorem: $\boxed{q(y) = \frac{1}{|a|} p_X\left(\frac{y-b}{a}\right)}$

Exp. $X \sim N(0,1)$

Q. $Y = aX + b$

$Y \sim ??$

Recall, $p_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

$p_Y(y) = \frac{1}{|a|} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-b)^2}{2a^2}}$

$= N(b, |a|^2)$

Obs. : $Y = \sigma X + \mu$

then $Y \sim N(\mu, \sigma^2)$

Cor. Recall the Gaussian

Φ - function

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2} \, dx$$

$$\begin{aligned}
N(\mu, \sigma^2)([a,b]) &= P\left(a \le N(\mu, \sigma^2) \le b\right) \\
&= P\left(a \le \sigma N(0,1) + \mu \le b\right) \\
&= P\left(\frac{a-\mu}{\sigma} \le N(0,1) \le \frac{b-\mu}{\sigma}\right) \\
&= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)
\end{aligned}$$

Remark. If $X \sim N(\mu, \sigma^2)$

Then the next result shows that X with high probability (more than 99.7%) attains the values in $[\mu - 3\sigma, \mu + 3\sigma]$.

[Therefore, one may assume X maps into $[\mu - 3\sigma, \mu + 3\sigma]$]

3σ - Rule

If $X \sim N(\mu, \sigma^2)$ then

$$P(|X - \mu| \leq 2\sigma) \geq 0.954$$

and $$P(|X - \mu| \leq 3\sigma) \geq 0.997$$

Pf.

$$P(|X - \mu| \leq \alpha\sigma)$$

$$= P(\mu - \alpha\sigma \leq X \leq \mu + \alpha\sigma)$$

$$= \Phi(\alpha) - X(-\alpha)$$

For $\alpha = 2$, $P(|X - \mu| \leq 2\sigma) = \Phi(2) - \Phi(-2) = 0.9545$

$\alpha = 3$, $P(|X - \mu| \leq 3\sigma) = \Phi(3) - \Phi(-3) = \cdot 9973$

Recall

Ω

Random vectors; $X_1, X_2, \ldots, X_n$

$$X_i : \Omega \to \mathbb{R}$$

$$\vec{X} = (X_1, \ldots X_n) : \Omega \to \mathbb{R}^n$$

$$\vec{X}(w) = (x_1, x_2, \ldots, x_n)$$

Prob. distribution.

$$P(\vec{X} \in B) = P_{(X_1, \ldots, X_n)}(B)$$

$$B \in \mathcal{B}(\mathbb{R}^n).$$

~~many~~

Joint density fn.

A random vector $\vec{X} = (X_1 \ldots X_n)$ is said to be continuous if there is a function $p : \mathbb{R}^n \to \mathbb{R}$ such that $a_j < b_j$, $1 \le j \le n$

$$P(a_1 \le X_1 \le b_1, \ldots, a_n \le X_n \le b_n) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} p(x_1 \ldots x_n)\, dx_n \ldots dx_1$$

marginal distribution.

Given $P_{\vec{X}}(x_1, \dots, x_n)$

and $B \in \mathcal{B}(\mathbb{R})$.

$$P_{X_j}(B) = P_{(X_1, \dots, X_n)}(\mathbb{R} \times \dots \times \mathbb{R} \times \underset{\substack{\downarrow \\ j\text{th one}}}{B} \times \mathbb{R} \dots \times \mathbb{R})$$

$$= P(-\infty < X_1 < \infty, \dots, -\infty < X_{j-1} < \infty,$$
$$a \leq X_j \leq b, \dots, -\infty < X_n < \infty)$$

is called the marginal distribution.

Proposition. If a random vector $\vec{X} = (X_1, \dots, X_n)$ has density $p : \mathbb{R}^n \to \mathbb{R}$ then for each $j \leq n$, the random variable X_j is continuous with density

$$p_j(x_j) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(\dots, x_{j-1}, x_j, x_{j+1} \dots) \, dx_n \dots dx_{j+1} \, dx_{j-1} \dots dx$$

In $n=2$

$$p_{(X_1,X_2)}(x_1,x_2) = \text{—}$$

$$p_{X_1}(x_1) = \int_{-\infty}^{\infty} p(x_1,x_2)\,dx_2$$

$$p_{X_2}(x_2) = \int_{-\infty}^{\infty} p(x_1,x_2)\,dx_1$$

Exp. Choose by random a point $x = (x_1, x_2, x_3)$ in the unit ball of $\mathbb{R}^3$.

How ~~Have~~ are the coordinates x_1, x_2, x_3 are distributed?

Pf. Let $\vec{X} = (X_1, X_2, X_3)$ be uniformly distributed on the unit ball $K = \{(x_1, x_2, x_3) \mid x_1^2+x_2^2+x_3^2 \le 1\}$

$$p_{\vec{X}}(x) = \begin{cases} \frac{3}{4\pi}, & x \in K \\ 0, & \text{otherwise} \end{cases}$$