2. Continuous Probability Distributions, Estimation using Normal Distribution

04 April 2024 12:26

CONTINUOUS PROBABILITY DISTRIBUTION

Probability density function

For every $x_i \in T_X$ (where X is continuous), we assign a real numbers $P(X)$ satisfying the conditions:

1. $P(X = x_i) \geq 0$
2. $\int_{support(x)} p(x_i)\, dx = 1$

$P(X)$ is called the pdf of X.

Cumulative distribution function

If X $\longrightarrow$ continuous random variable with pdf $p(x)$, then the function $F(x)$ is defined by

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} p(x)\, dx$$

Mean and variance of continuous random variables

$$\mu = \int_{-\infty}^{\infty} x \cdot p(x)$$

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot p(x)$$

NORMAL DISTRIBUTION

The continuous probability distribution having the pdf $p(x)$ given by:

$$\boxed{p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}}$$

$-\infty < x < \infty$

$-\infty < \mu < \infty$

$\sigma > 0$

The graph of $p(x)$ $\longrightarrow$ bell curve

symmetric about the line $x = \mu$

$\longrightarrow$ Normal probability curve

STANDARDISATION

$X \sim \text{Normal}(\mu, \sigma^2)$

We have some $Z = (X-\mu/\sigma) \sim \text{Normal}(0,1)$

$$\text{PDF} = f_z(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}$$

$$\text{CDF} = F_z(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{\left(\frac{-u^2}{2}\right)} du$$

Say we need to find $P(a \le x \le b)$

$$P(a \le x \le b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx$$

This is complicated, so we use the following transformation:

$$\frac{x-\mu}{\sigma} = z \longrightarrow \text{standard normal variate } Z$$

$$x = z\sigma + \mu$$

$$dx = \sigma\, dz$$

$$\Rightarrow P(a \le x \le b) = \int_{z_1 = \frac{a-\mu}{\sigma}}^{z_2 = \frac{b-\mu}{\sigma}} \underbrace{\frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}}_{\text{standard normal}} dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{\frac{-z^2}{2}} dz$$

This is much easier for calculation.

EXPONENTIAL DISTRIBUTION

Let λ be a real constant > 0. Then the cdf of exponential distribution is given by:

$$\boxed{\begin{matrix} P(x) \\ \mathcal{E}(\lambda, x) \end{matrix} = \begin{cases} \lambda e^{-\lambda x} & , x \ge 0 \\ 0 & , x < 0 \end{cases}}$$

The cdf of the same is given by

$$\boxed{P(X \le a) = F_X(a) = 1 - e^{-\lambda a}}$$

Conditions

- $P(X) \ge 0$
- $\int_{-\infty}^{\infty} p(x)\,dx = 1$

Mean, variance, standard deviation

$$\boxed{\begin{aligned} E(X) &= \frac{1}{\lambda} \\ V(X) &= \frac{1}{\lambda^2} \\ \sigma &= \frac{1}{\lambda} \end{aligned}}$$

NORMAL APPROXIMATION TO BINOMIAL DISTRIBUTION

If $X \to$ Binomial (n, p), then

$$Z = \frac{X-\mu}{\sigma} = \frac{X-np}{\sqrt{np(1-p)}}$$

is approximately a standard normal random variable.

The probabilities involving X can be approximated by using a standard normal distribution.

The approximation is good $\iff$ n is large

For a good approximation,

$np > 5$ [OR] $n(1-p) > 5$

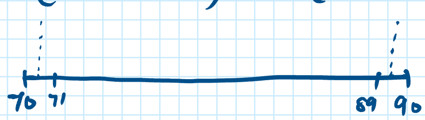
To approximate a binomial probability with a normal distribution, a continuity correction is applied as follows:

$$P(X \le x) \to P(X \le x+0.5) \cong P\left(Z \le \frac{x+0.5-np}{\sqrt{np(1-p)}}\right)$$

$$P(X \ge x) \to P(X \ge x-0.5) \cong P\left(Z \ge \frac{x-0.5-np}{\sqrt{np(1-p)}}\right)$$

Example

$$P(70 < X < 90) = P(71 \le X \le 89) \longrightarrow P(70.5 \le X \le 89.5)$$



NORMAL APPROXIMATION TO POISSON DISTRIBUTION

$X \to$ Poisson $(\lambda = np)$

$$Z = \frac{X-\mu}{\sigma} = \frac{X-\lambda}{\sqrt{\lambda}} = \frac{X-np}{\sqrt{np}}$$ $\longrightarrow$ approximately standard normal random variable

The approximation is good if

$\lambda > 5 \implies np > 5$

Again a continuity correction is applied:

$$P(X \le x) \longrightarrow P(X \le x+0.5)$$

$$P(X < x) = P(X \le x-1) \longrightarrow P(X \le x-0.5)$$

$$P(X \ge x) \longrightarrow P(X \ge x-0.5)$$

$$P(X > x) = P(X \ge x+1) \longrightarrow P(X \ge x+0.5)$$

$$P(x_1 < X < x_2) = P(x_1+1 \le X \le x_2-1) \longrightarrow P(x_1+0.5 \le X \le x_2-0.5)$$

$$P(x_1 < X < x_2) = P(x_1+1 \le X \le x_2-1) \longrightarrow P(x_1+0.5 \le X \le x_2-0.5)$$