Continuous Random Variables (II)

Goal: Study the Normal Distribution → useful in modelling measurements

Normal Distribution

Defined using two parameters:
- μ ⟶ mean
- σ^2 ⟶ variance

Denoted as $N(\mu, \sigma^2)$

Note: $\sigma = \sqrt{\sigma^2}$ is the standard deviation.

Suppose X is normally distributed ⟶ ie $X \sim N(\mu, \sigma^2)$

the pdf of X is defined as

$$\boxed{f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2} \qquad x \in (-\infty, \infty)}$$

To calculate the cdf of X,

$$\boxed{F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\,dt = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{x} e^{-(t-\mu)^2/2\sigma^2}\,dt}$$

$$\therefore \boxed{P(a \le X \le b) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_a^b e^{-(t-\mu)^2/2\sigma^2}\,dt}$$ ⟶ Not an easy integral to calculate!

Note:

① If $X \sim N(\mu, \sigma^2)$, then

$$\boxed{E(X) = \mu \quad \text{and} \quad V(X) = \sigma^2}$$

② When $\mu = 0$ and $\sigma^2 = 1$ ⟶ Standard Normal Dist → $Z \sim N(0,1)$

If $Z \sim N(0,1)$, the standard normal distribution

pdf: $$\boxed{f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \qquad z \in (-\infty, \infty)}$$

To be specific about parameters

denote pdf of $X \sim N(\mu, \sigma^2)$ ⟶ $f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2} \quad x \in (-\infty, \infty)$

If $X \sim N(\mu, \sigma^2)$ and $Z \sim N(0,1)$

Change of variable with $s = \frac{(t-\mu)}{\sigma}$

$$P(X \le x) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{x} e^{-(t-\mu)^2/2\sigma^2}\,dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{(x-\mu)}{\sigma}} e^{-s^2/2}\,ds$$

$$= P\left(Z \le \frac{(x-\mu)}{\sigma}\right)$$

∴ If $X \sim N(\mu, \sigma^2)$ then $\frac{X-\mu}{\sigma}$ has the standard normal distribution.

If $Z \sim N(0,1)$ ⟶ cdf of Z is denoted as $\Phi(z)$

ie $$\boxed{P(Z \le z) = \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\,dt}$$

③ If $X \sim N(\mu, \sigma^2)$ then $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

$$\therefore P(a \le X \le b) = P\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right)$$

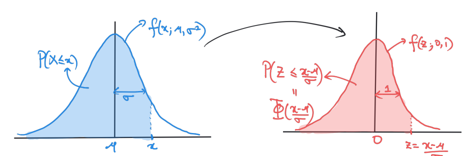
$$= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

∴ Can use the standard normal cdf to calculate probabilities for the random variable $X \sim N(\mu, \sigma^2)$

$X \sim N(\mu, \sigma^2) \longrightarrow Z \sim N(0,1)$

$x \longmapsto z = \frac{x-\mu}{\sigma}$ ⟶ "z-score" of x

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



Percentiles and Critical Values

Suppose $Z \sim N(0,1)$ ⟶ standard normal distribution

For $\alpha \in (0,1)$, z_α ⟶ α^{th} critical value satisfies

$$\boxed{P(Z \ge z_\alpha) = \alpha} \quad \text{ie} \quad \boxed{\int_{z_\alpha}^{\infty} f(z;0,1)\,dz = \alpha}$$

ie $1 - P(Z \le z_\alpha) = \alpha \Rightarrow 1 - \Phi(z_\alpha) = \alpha$

Let $p \in (0,1)$ ⟶ $(100p)^{th}$ percentile, $\eta(p)$ satisfies

$$\boxed{P(Z \le \eta(p)) = p}$$

ie $\Phi(\eta(p)) = p$

Suppose $X \sim N(\mu, \sigma^2)$ and $\alpha \in (0,1)$

$$\boxed{\alpha^{th} \text{ critical value for } X := x_\alpha = \sigma \cdot z_\alpha + \mu}$$

$z_\alpha = \alpha^{th}$ critical value for $Z \sim N(0,1)$.

Approximating Bin(n,p) using Normal Distribution

Let $X \sim Bin(n,p)$ ie $\mu_X = np$ and $\sigma_X^2 = n \cdot p \cdot q$

If the $Bin(n,p)$ is not too skewed

$$\boxed{P(X \le x) \approx \Phi\left(\frac{x + 0.5 - \mu_X}{\sigma_X}\right)}$$

The right hand side does not involve terms of the form $\binom{n}{k}$ ⟶ which can be difficult to compute.