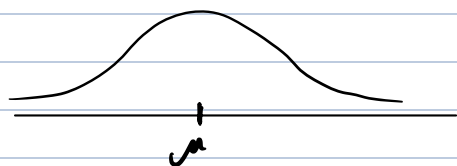


Lecture 11 Part 2: Normal distribution

DEF: If X is a normal r.v. with parameters μ and σ^2 , then the pdf of X is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad \text{for } x \in \mathbb{R}$$


* Comes up in Central Limit Theorem

Note: The normal distribution is also called the Gaussian distribution

Facts about the normal distribution $X \sim N(\mu, \sigma^2)$

- Mgf: $M_X(t) = \mathbb{E}[e^{tx}] = e^{\mu t + \sigma^2 t^2 / 2}$

Proof: Exercise

$$\mathbb{E}[e^{tx}] = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}(x-\mu)^2 + tx} dx$$

Idea: Complete the square

- Moments: $\mathbb{E}[X] = \mu$ and $\text{var}(X) = \sigma^2$

Proof: $\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx$

less annoying: use mgf

$$\mathbb{E}[X] = \left. \frac{d}{dt} M_X(t) \right|_{t=0} = \left. (\underbrace{\mu}_{=0} + \underbrace{\sigma^2 t}_{=1}) e^{\mu t + \sigma^2 t^2 / 2} \right|_{t=0} = \mu$$

$$\text{var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}[X^2] - \mu^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2 \checkmark$$

$$\mathbb{E}[X^2] = \left. \frac{d^2}{dt^2} M_X(t) \right|_{t=0} = \dots = \sigma^2 + \mu^2$$

exercise

Linear combinations of Normal r.v.s

Suppose $X \sim N(\mu_1, \sigma_1^2)$ and that $Y \sim N(\mu_2, \sigma_2^2)$, X and Y are independent

① $aX + b \sim N(a\mu_1 + b, a^2\sigma_1^2)$

Proof: Want to calculate $f_Y(y)$ where $Y = aX + b$

First, calculate $F_Y(y)$

$$\begin{aligned} F_Y(y) &= \mathbb{P}(aX + b \leq y) \\ &= \mathbb{P}\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right) \end{aligned}$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X\left(\frac{y-b}{a}\right) = f_X\left(\frac{y-b}{a}\right) \cdot \frac{d}{dy} \left(\frac{y-b}{a}\right) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)$$

Recall that $f_X(x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{1}{2\sigma_1^2}(x-\mu_1)^2}$

Exercise: show that $\frac{1}{a} f_X\left(\frac{y-b}{a}\right)$ corresponds to the pdf of a $N(a\mu_1 + b, a^2\sigma_1^2)$

② $(a_1 X_1 + b_1) + (a_2 Y + b_2) \sim N(\quad)$ (do this when introducing bivariate normal)

Standard Normal r.v.

If $Z \sim N(0, 1)$ Then we say "Z follows the standard normal distribution" or we say "Z is a standard normal r.v." Z has pdf $f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$

Fact (uses linear transformation of r.v.) : If $X \sim N(\mu, \sigma^2)$, then $\frac{X-\mu}{\sigma} \stackrel{d}{=} Z$ when $Z \sim N(0, 1)$

Exercise : • show that $E[Z] = 0$ and $\text{var}(Z) = 1$ using mgfs
• show that $E\left[\frac{X-\mu}{\sigma}\right] = 0$ and $\text{var}\left(\frac{X-\mu}{\sigma}\right) = 1$ use mgfs
• show that $\frac{X-\mu}{\sigma} \stackrel{d}{=} Z$ using mgfs

Need to show that mgf of $\frac{X-\mu}{\sigma}$ is equal to mgf of Z

Another useful fact : Suppose we want to calculate $F_X(x) = P(X \leq x)$
$$P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = P\left(Z \leq \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

 $Z \sim N(0, 1)$ ↑ symbol for F_Z

Exercise : Suppose $X \sim N(3, 9)$, Find

(a) $P(2 < X < 5) = \Phi\left(\frac{2}{3}\right) - \Phi\left(-\frac{1}{3}\right)$

$$P\left(\frac{2-3}{3} < \frac{X-3}{3} < \frac{5-3}{3}\right) = P\left(-\frac{1}{3} < Z < \frac{2}{3}\right) = \Phi\left(\frac{2}{3}\right) - \Phi\left(-\frac{1}{3}\right)$$

(b) $P(X > 0) = P\left(\frac{X-3}{3} > -1\right) = 1 - \Phi(-1)$

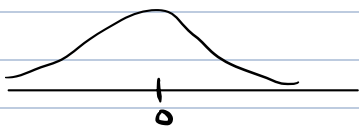
(c) $P(|X-3| > 6) = P\left(\frac{|X-3|}{3} > 2\right)$

$$\begin{aligned} &= P(|Z| > 2) = P(Z > 2 \text{ or } Z < -2) \\ &= P(Z > 2) + P(Z < -2) \\ &= P(Z > 2) + 1 - P(Z \leq 2) \\ &= 2(1 - \Phi(2)) \end{aligned}$$

Fact about Standard Normals : let $z \geq 0$.

Then $\Phi(-z) = P(Z \leq -z) = P(-Z \geq z) = P(Z \geq z) = 1 - P(Z \leq z) = 1 - \Phi(z)$

Since Z is symmetric around 0, we also have that $Z \stackrel{d}{=} -Z$



Bivariate and Multivariate normal distributions with mean μ and covariance Σ

DEF: If (X_1, X_2) have a bivariate normal distribution, the pdf is given by the following: Define $X = (X_1, X_2)$

$$f_X(x) = (2\pi)^{-1} \det(\Sigma)^{-1/2} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}$$

$\uparrow$
 (x_1, x_2)



$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho \\ \rho & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} \text{Var}(X_1) & \text{Corr}(X_1, X_2) \\ \text{Corr}(X_1, X_2) & \text{Var}(X_2) \end{pmatrix}$$

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} \mathbb{E}[X_1] \\ \mathbb{E}[X_2] \end{pmatrix}$$

Univariate normal: mean $\mu \in \mathbb{R}$

Bivariate normal: mean $\mu \in \mathbb{R}^2$

d-dimensional multivariate normal: mean $\mu \in \mathbb{R}^d$

variance $\sigma^2 > 0$

covariance matrix $\Sigma \in \mathbb{R}^{2 \times 2}$

(positive definite)

covariance matrix $\Sigma \in \mathbb{R}^{d \times d}$

(positive definite)

DEF: $\Sigma \in \mathbb{R}^{d \times d}$ is positive definite if for all $x \in \mathbb{R}^d$, we have

$$x^T \Sigma x > 0$$

Some facts

① Suppose $(X_1, X_2) \sim N(\mu, \Sigma)$. Then $X_1 \perp\!\!\!\perp X_2 \Leftrightarrow \text{Corr}(X_1, X_2) = \Sigma_{12} = \Sigma_{21} = 0$

(Remember $\perp\!\!\!\perp \Rightarrow \text{Corr} = 0$

$\text{Corr} = 0 \not\Rightarrow \perp\!\!\!\perp$ in general)