5.4 (General) Normal Distribution

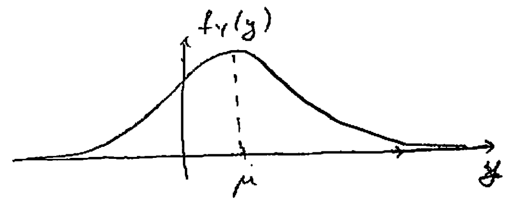
Ex Let $X \sim N(0,1)$, $\mu \in \mathbb{R}$, $\sigma > 0$.
Compute the pdf of $Y = \mu + \sigma X$.

Apply Thm on transformations (p. 53) with $y = g(x) = \mu + \sigma x$. (linear transformation)

$$f_Y(y) = f_X(x) \cdot \Big|\underbrace{\frac{d}{dx} g(x)}_{=\sigma}\Big|^{-1} = \frac{1}{\sigma} f_X\left(\frac{y-\mu}{\sigma}\right) \quad (*)$$

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

$$= \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/2\sigma^2}.$$



- Remark: The first part of this computation, (*), gives the pdf of any linear transformation of any r.v.; if $Y = \mu + \sigma X$ then $f_Y(y) = \frac{1}{|\sigma|} f_X\left(\frac{y-\mu}{\sigma}\right)$

Ex Compute $E[Y]$ and $Var(Y)$, σ_Y

$E[Y] = E[\mu + \sigma X] = \mu + \sigma \underbrace{E[X]}_{=0}$ (by properties of expectation) (since $X \sim N(0,1)$)

$E(Y) = \mu$.

$Var(Y) = Var(\mu + \sigma X) = Var(\sigma X) = \sigma^2 \underbrace{Var(X)}_{=1}$ (by properties of variance) (since $X \sim N(0,1)$)

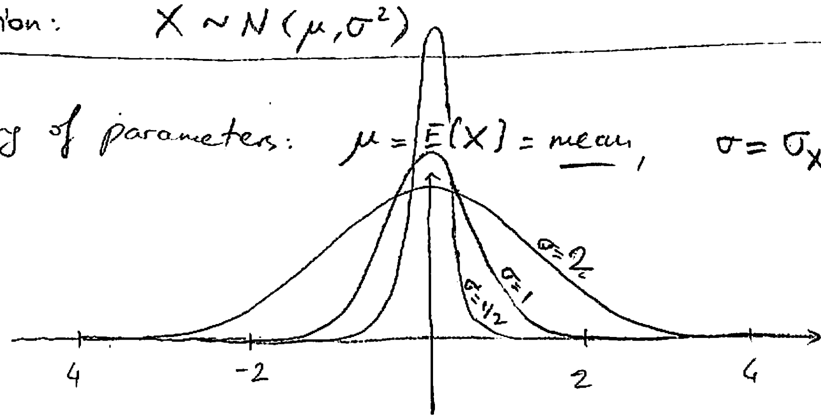
$Var(Y) = \sigma^2$, $\Rightarrow$ $\sigma_Y = \sigma$.

Def (Normal distribution) A r.v X has normal distribution with parameters μ and σ if X has pdf

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty.$$

Notation: $X \sim N(\mu, \sigma^2)$

- Meaning of parameters: $\mu = E[X] =$ mean, $\sigma = \sigma_X =$ st. deviation of X



Remark (Z-score) Let X be a r.v. with $E[X]=\mu$, $Var(X)=\sigma^2$.

Consider the r.v. $Z = \dfrac{X-\mu}{\sigma}$.

Then $E[Z]=0$, $Var(Z)=1$ (Check as in Ex. p. 58!)

Z is called the "Z-score" or the "standard score" of X.

Example: $X \sim N(\mu,\sigma) \Rightarrow Z \sim N(0,1)$

Advantage of Z: (a) no units, (b) pdf does not depend on any parameters.

Ex The GRE scores are normally distributed with mean 500 and st. dev. 100. What score would place a student in the top 10% ?

$X \sim N(500, 100^2)$. Looking for x: $P\{X \le x\} = 0.9$

$$Z = \frac{X-500}{100} \sim N(0,1)$$

$$P\{X \le x\} = P\Big\{\underbrace{\frac{X-500}{100}}_{= Z} \le \frac{x-500}{100}\Big\} = \Phi\left(\frac{x-500}{100}\right)$$

Now solve the equation $\Phi\left(\frac{x-500}{100}\right) = 0.9$

The table (p. 201 of Ross) gives

$$\frac{x-500}{100} \approx 1.28 \Rightarrow x = 628$$

Ex A manufacturer claims that 99% of the light bulbs it makes last at least 900 hrs. A test reveals that the lifetime of a bulb is normally distributed with mean 1000 hrs, st. dev. 50 hrs. Is the manufacturer's claim correct?

X = lifetime of a bulb $\sim N(1000, 50^2)$ Claim: "$P\{X \le 900\} \le 0.01$"

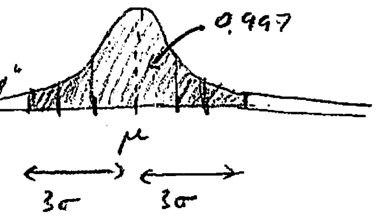
Reality: $P\{X \le 900\} = P\Big\{\underbrace{\frac{X-1000}{50}}_{= Z \sim N(0,1)} \le \underbrace{\frac{900-1000}{50}}_{= -2}\Big\} = \Phi(-2) = 1 - \Phi(2)$ (by (*) p. 57)

$\approx 1 - 0.977 = 0.023. = 2.3\%$

Therefore, the claim is false.

Ex $X \sim N(\mu, \sigma^2)$. Then

$$\boxed{P\{|X-\mu| \le 3\sigma\} \approx 0.997.}$$ "Deviation inequality"



Indeed, $Z := \frac{X-\mu}{\sigma} \sim N(0,1)$.

$$P\{|X-\mu| \le 3\sigma\} = P\{|Z| \le 3\} \approx 0.997.$$

In words: { "A normal random variable stays within 3 standard deviations from its mean with probability 99.7%". }

Lec. 22 03/04

5.4.1 Normal approximation to Binomial

- Recall Poisson approx. (4.7), which is valid for rare successes ($p \to 0$):

$$\text{Binom}(n,p) \approx \text{Poisson}(\lambda), \qquad \text{if } n \to \infty \text{ and } np \to \lambda = \text{const}.$$

- Normal approx. holds when successes are not rare ($p = \text{const}$):

$$\boxed{\text{Binom}(n,p) \approx N(\mu, \sigma^2) \text{ where } \mu = np,\ \sigma^2 = np(1-p), \text{ if } n \to \infty \text{ and } p = \text{const}}$$

In other words, if $X \sim \text{Binom}(n,p)$, its Z-score $Z = \frac{X-np}{\sqrt{np(1-p)}}$ is approximately $N(0,1)$

- Formally, this is the content of the Central Limit Theorem (CLT):

Thm (De Moivre-Laplace CLT):

> Let $X \sim \text{Binom}(n,p)$; consider the Z-score $Z = \frac{X-np}{\sqrt{np(1-p)}}$.
>
> Then, for every $a \in \mathbb{R}$,
>
> $$\underbrace{F_Z(a)}_{\text{cdf of } Z} \longrightarrow \underbrace{\Phi(a)}_{\text{cdf of } N(0,1)} \quad \text{as } n \to \infty,\ p \text{ fixed}$$

In particular, $P\{a \le Z \le b\} \to \Phi(b) - \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}\, dx$, as $n \to \infty$, p fixed

$$\| $$

$$P\left\{a \le \frac{X-np}{\sqrt{np(1-p)}} \le b\right\}$$