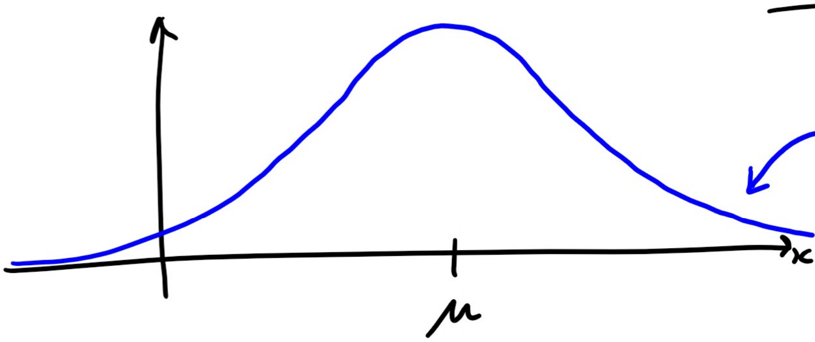
3Y03 - 3J04 PROBABILITY & STATISTICS FOR (CIVIL) ENGINEERING

C01 - Lecture 13

Last time — NORMAL DISTRIBUTION



- symmetric, bell curve

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

$(\mu \in \mathbb{R}, \sigma > 0)$.

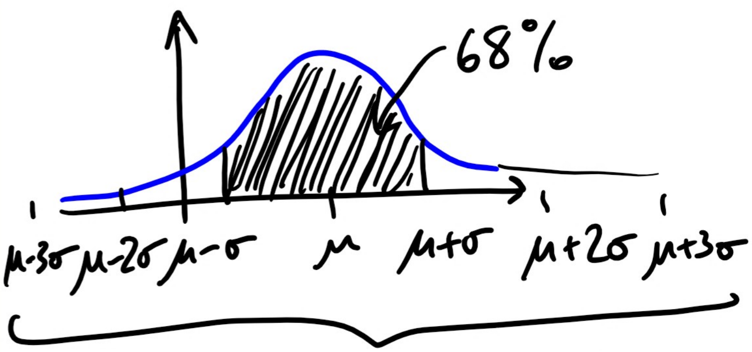
$$P(X < \mu) = P(X > \mu) = 0.5.$$

Working out $P(X < x) = F(x)$ i.e. integrating $f(x)$ not possible (no closed form expression).

↳ need numerical computation techniques

For any normal distribution mean μ, variance σ^2

$X \sim N(\mu, \sigma^2)$ we have a "68-95-99.7" rule

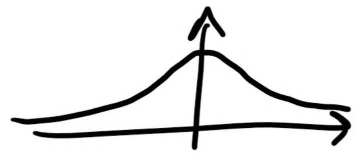


$$P(\mu - \sigma < X < \mu + \sigma) = 0.68$$
$$P(\mu - 2\sigma < X < \mu + 2\sigma) = 0.95$$
$$P(\mu - 3\sigma < X < \mu + 3\sigma) = 0.997$$

Whole span $\approx 6\sigma$

There is a standard normal distribution

$$Z \sim N(0, 1)$$



We denote $P(Z \leq z) = \Phi(z)$ (c.d.f.)

It would be impractical to have a table for calculating $P(X \leq x)$ for every $X \sim N(\mu, \sigma^2)$. Instead we have one table, for $Z \sim N(0, 1)$ & <u>standardize</u> any other $X \sim N(\mu, \sigma^2)$:

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

$$P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right)$$

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

$$= P\left(Z \leq \underbrace{\frac{x-\mu}{\sigma}}_{z\text{-value}}\right)$$

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

This is NOT an approximation. These are equivalent statements (one about X, the other about Z).

<u>Example</u> File transfer speed from server to computer is normally distributed with mean 5.75 Mbps and variance 0.35^2 $(\text{Mbps})^2$

(a) Find probability that speed is 6.7 Mbps or more.

(b) " " " " " 5.5 Mbps or slower.

(c) Find speed such that 90% of the time transfer is faster than that speed.

(d) Find symmetric interval of speed such that 99% of the time transfer speed is in that interval

Solution X = transfer speed $\sim N(5.75, 0.35^2)$.

(a) $P(X \geqslant 6.7)$ (b) $P(X \leq 5.5)$

(c) If $P(X \geqslant s) = 0.9$, then $s = ?$

(d) If $(\mu - t, \mu + t)$ with $P(\mu - t \leq X \leq \mu + t) = 0.99$ then $t = ?$

doesn't matter for continuous r.v. X <

(a)
$$P(X \geqslant 6.7) = 1 - P(X \leq 6.7)$$
$$= 1 - P\left(Z \leq \frac{6.7 - \mu}{\sigma}\right)$$
$$= 1 - P\left(Z \leq \frac{6.7 - 5.75}{0.35}\right)$$
$$= 1 - P(Z \leq 2.71)$$ ← 2 d.p.

use table

$$= 1 - 0.996636$$
$$= \underline{0.003364}.$$

(b)
$$P(X \leq 5.5) = P\left(Z \leq \frac{5.5 - \mu}{\sigma}\right)$$
$$= P\left(Z \leq \frac{5.5 - 5.75}{0.35}\right) = P(Z \leq -0.71)$$

2dp.

$$= \underbrace{0.238852}.$$ ↙ use table

(c) If $P(X \geq s) = 0.9$, what is s?

$$P(X \leq s) = 0.1$$

$$\text{i.e. } P\Big(Z \leq \underbrace{\frac{s-\mu}{\sigma}}_{\approx -1.28}\Big) = 0.1$$

≈ -1.28 from table

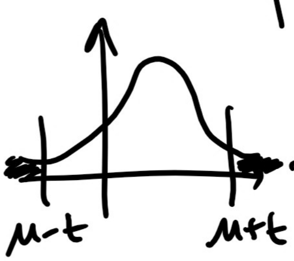
This is an approximation of our z-value as there is no value in the table giving probability exactly 0.1 (the closest value to 0.1 is given by −1.28).

Solve for s: $s \cong (-1.28)(0.35) + 5.75$

$s \cong 5.302.$

(d) $P(\mu - t \leq X \leq \mu + t) = 0.99$

$$P(X \leq \mu - t) + P(X \geq \mu + t) = 1 - 0.99 = 0.01$$



2 equal tail ends

$$\text{i.e. } P(X \leq \mu - t) = \frac{0.01}{2} = 0.005.$$

$$\text{i.e. } P\Big(Z \leq \frac{\mu - t - \mu}{\sigma}\Big) = 0.005$$

↳ look in table

$$\Big(= P\Big(Z \leq -\frac{t}{\sigma}\Big)\Big)$$

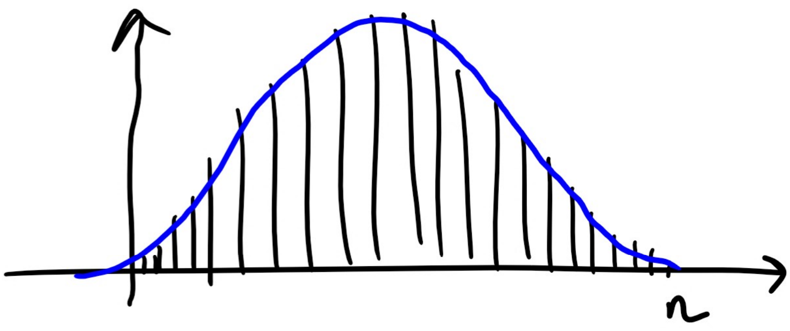
↳ ≈ -2.58 → solve for t

Gives the value in the table closest to 0.005.

& get $t = 0.903$

So interval: $(\mu - t, \mu + t) = \underline{(4.847, 6.653)}.$

We said normal distribution is long-term approximation of many trials — indeed "Normal approximates Binomial (n,p) for large n."



If $X \sim Bin(n,p)$ then recall it has $\mu = np$, $\sigma^2 = np(1-p)$

We approximate X by $N(np, np(1-p))$

In other words $Z = \dfrac{X - np}{\sqrt{np(1-p)}} \cancel{\sim} N(0,1)$.

better said $N(0,1)$ is a good approx. of Z so defined here.

But we can do better with a continuity correction:

$$P(X \leq x) = P(X \leq x + 0.5)$$
$$= P\left(Z \leq \frac{x + 0.5 - np}{\sqrt{np(1-p)}}\right).$$

To Be Continued ...