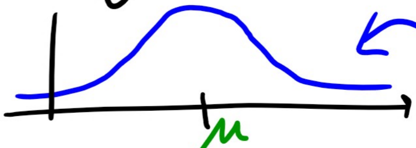
3Y03 – Probability and Statistics for Engineering

WS19 **Lecture 14**

Last Time

Normal Distribution

given in terms of its mean μ & variance σ^2 i.e. $X \sim N(\mu, \sigma^2)$.



bell-shaped curve (symmetric about $x = \mu$)

want (as always) $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx$ BUT cannot integrate p.d.f. $f(x)$

So standardize: $Z := \dfrac{X - \mu}{\sigma} \sim N(0,1)$ and

$$F(x) = P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \underbrace{\frac{x-\mu}{\sigma}}_{=:z}\right) = P(Z \leq z) = \boxed{\Phi(z)}$$

c.d.f.

(LOOK UP IN TABLE.)

Example File transfer speed X from server to computer

$X \sim N(5.75, 0.35^2)$

last time: (b) $P(X \leq 5.5) = P\left(Z \leq \dfrac{5.5 - 5.75}{0.35}\right) = P(Z \leq \underbrace{-0.71}_{\substack{-0.7 \\ -0.01}})$

$= 0.238852.$

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

i.e. Find s s.t. $P(X > s) = 0.9$.

" $P\left(Z > \dfrac{s - 5.75}{0.35}\right) = \underbrace{0.9}$

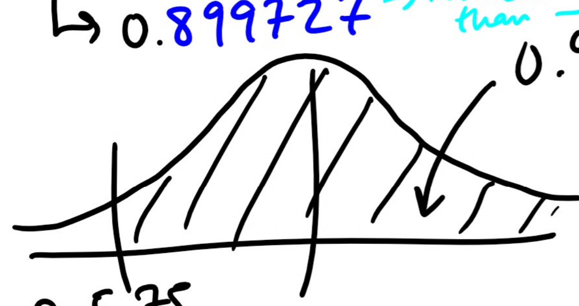
Find in table

For our purposes enough to choose whichever # (to 2d.p.) gives closest answer

Here : choice between $\underline{\underline{1.28}}$ & $1.29 \rightarrow 0.901475$

$\hookrightarrow 0.899727 \rightarrow$ is closer to 0.9 than

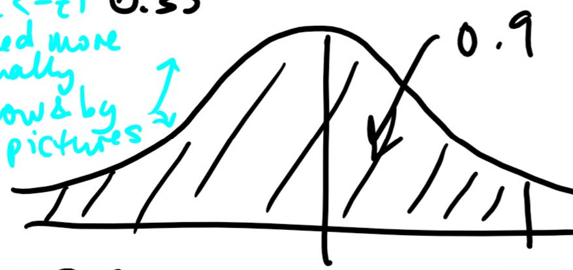
$$0.9 = 1 - \underbrace{P\left(Z < \frac{s-5.75}{0.35}\right)}_{=\,0.1}$$



Either look up 0.1 instead of 0.9 in table & find z-value −1.28

or use $P(Z<z) = 1-P(Z<-z)$ explained more formally below & by the pictures

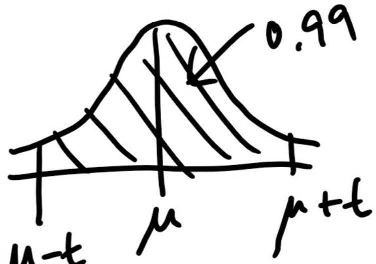
$$\text{So} \quad 1.28 = -\left(\frac{s-5.75}{0.35}\right)$$



$$\text{So} \quad s = -(1.28)(0.35) + 5.75$$
$$= 5.302.$$

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

i.e. Find t s.t. $P(5.75 - t < X < 5.75 + t) = 0.99$



(interval : $(\mu - t, \mu + t)$)

$P(a<X<b) = P(X<b) - P(X<a)$

$$= P(X < 5.75 + t)$$
$$- P(X \cancel{\geq} < 5.75 - t)$$

Sorry! Mistake!

$$= (1 - P(X > 5.75 + t)) - P(X < 5.75 - t)$$
$$= 1 - (P(X < 5.75 - t) + P(X > 5.75 - t))$$

2 tails

Tails ($P(X<5.75-t)$ and $P(X>5.75+t)$) must be equal to each other & equal to

$$\frac{1-0.99}{2} = 0.005$$

So solve for t in $P(X < 5.75 - t) = 0.005$

(Choose one.)

$$= P\left(Z < \frac{(5.75 - t) - 5.75}{0.35}\right)$$

$$= P\left(Z < {-t}/{0.35}\right)$$

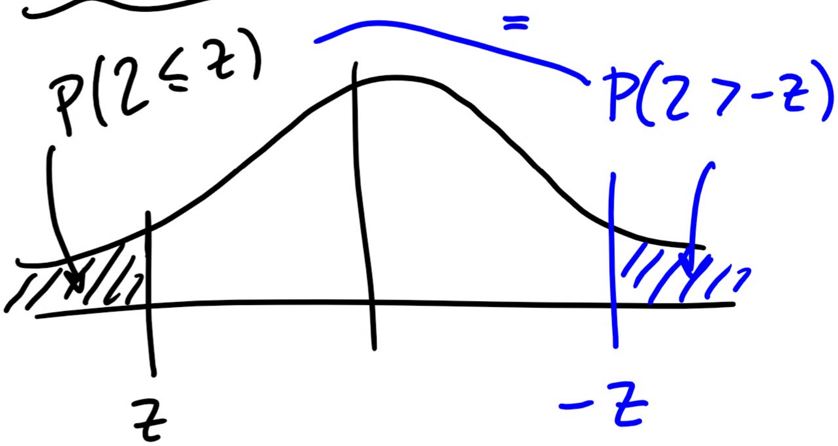
Look up in table.
Get $-2.\cancel{8}5\,58$

So $$\frac{-t}{0.35} = -2.\cancel{85}^{\,58}$$

$$\Rightarrow t = 0.903.$$

So interval: $(5.75 - 0.903,\ 5.75 + 0.903)$

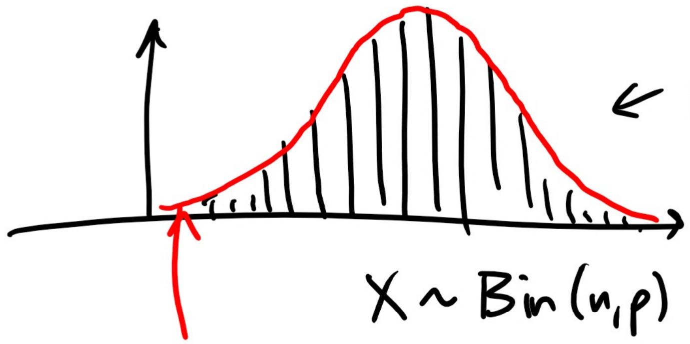
$= (4.847,\ 6.653).$



If you only have positive (RHS) part of Standard Normal Table:

$$P(Z \le z) = P(Z \ge -z)$$
$$= 1 - P(Z \le -z).$$

Normal Distribution to approximate Binomial Distribution



$X \sim Bin(n,p)$ $\qquad f(x) = \binom{n}{x} p^x (1-p)^{n-x}$

We approximate Bin(n,p) with Normal distribution

$N(np,\ np(1-p))$ i.e.

If $X \sim Bin(n,p)$ then $Z = \dfrac{X - np}{\sqrt{np(1-p)}}$ is approximated by $N(0,1)$

To improve the approximation we do a "continuity correction":

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

Similarly: $P(X \geq x) \approx P\left(Z \geq \dfrac{x - 0.5 - np}{\sqrt{np(1-p)}}\right)$

$$P(x_1 \leq X \leq x_2) \approx P\left(\frac{x_1 - 0.5 - np}{\sqrt{np(1-p)}} \leq Z \leq \frac{x_2 + 0.5 - np}{\sqrt{np(1-p)}}\right)$$

(Make region bigger.)

In this context this is a valid approximation

if $\boxed{\begin{matrix} np > 5 \\ n(1-p) > 5 \end{matrix}}$

n "large" means that the expected # of successes & expected # of failures are both > 5.

Example Multiple Choice Test with 60 questions 5 choices per question. Guess randomly with equal prob. ...

Find prob. of getting between 10 & 20 correct (inclusive).

Solution X = # correct answers $\sim \text{Bin}(60, \frac{1}{5})$

So $$P(10 \leq X \leq 20) = \sum_{x=10}^{20} \binom{60}{x} \left(\frac{1}{5}\right)^x \left(\frac{4}{5}\right)^{60-x}$$

This sum is not very nice to compute! Easier to use Normal approx.

Check : $np = 60\left(\frac{1}{5}\right) = 12 > 5$

Mean

$n(1-p) = 48 > 5$ Use Normal approx.

$$P(10 \leq X \leq 20) \overset{\text{cont. correction}}{=} P(9.5 \leq X \leq 20.5)$$

$$\cong P\left(\frac{9.5 - 12}{\sqrt{9.6}} \leq Z \leq \frac{20.5 - 12}{\sqrt{9.6}}\right)$$

Variance :

$np(1-p) = 60\left(\frac{1}{5}\right)\left(\frac{4}{5}\right)$

$= 9.6$

$\vdots$ Check !

$P(-0.81 \leq Z \leq 2.74)$

$= P(Z \leq 2.74) - P(Z \leq -0.81)$

$= 0.996928 - 0.208970$

$= 0.787958$

$= 0.788.$

Using the Binomial (sum above) actual answer

$= 0.782.$ 0.7819595