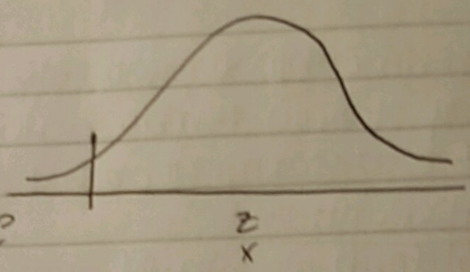
Lecture 3

03/20/19



The Normal Distribution

- can do this in excel
- Find μ and σ
- using the normal distribution
- what's a score that will cut the pop?
  - → what is the cut off - point?
  - "inverse normal distribution"
  - (slide 24/27 session 2)
- inverse normal distribution on excel
- "STANDARDIZATION"
  - → normal, as well as other variables for other analysis
  - if $z = 1.28$, find x on the standard normal
    - $x = 71.26$
  - what score corresponds ~~to~~ to that probability?

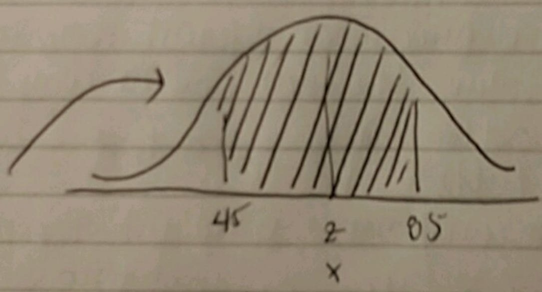
• examples:

$\bar{x} = 54.1$

$s = 13.41$

P(less than 30)?

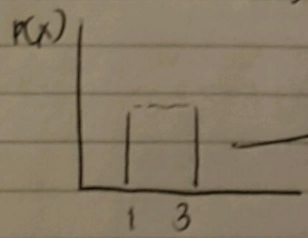
P(45 and 85)?



P(85) - P(45)

Sampling Distribution

- inferential statistics ⟶ super important
  - → use statistics to generalise
- • keep sampling means ⟶ create another distribution
- • small size vs sample size



⟶ all have the same probabilities in the uniform distribution

- eg. mean of means ⟶ unbiased estimator of population mean
- proportion to the sample size

Lecture 3

- Inferential statistics

- Basic Assumptions:
  - $Y \sim N(\mu, \sigma^2)$ → the variable we're observing is normal
  - observations are independent and identically distributed
    - → this is the assumption that's more commonly violated
    - → independence
    - → conditional on the mean
- Inference when σ is known
  - the z test
  - calculate the cutoff ⟶ find the upper & lower limits
  - probability of observing $\bar{x} = 100$?
  - $\sigma_{\bar{x}}$ ⟶ needs to be SD of sampling rather than original
    - ↳ $\frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{25}}$
  - p (observing 3.33 or less) ⟶ but want the prob to the right, so $1 - P(x)$
  - depends on the sample size
  - potential sample size
  - confidence ~~intervals~~ intervals
    - upper limit
    - lower limit
    - willingness to make a Type 1 error
    - corresponds to the cutoff value that we get when
    - 99% correct of 99% of the time ⟶ enlarge confidence
  - confidence intervals for future forecasts intervals
  - probability of observing mean = 0.95 ⟶ can't say this!

- Inference when σ is unknown
  - one sample t-test
  - estimate σ from the sample, since we don't know σ of populati
    - ↳ But "penalize" ourselves for the fact that we don't know the σ from the sample
    - ways:
      1) Having a tail
      2) use degrees of freedom
  - σ → calculate sum of squares
- do the same thing in all these cases generally

TEXTBOOK
TA | PROF perpert
BOOK list SD
review Economy
3 weeks time

$$\rightarrow \boxed{Z = \frac{x - \mu}{\sigma_x}} = \frac{\text{observed (sample)} - \text{expected (pop)}}{SD_{\text{observed}}}$$

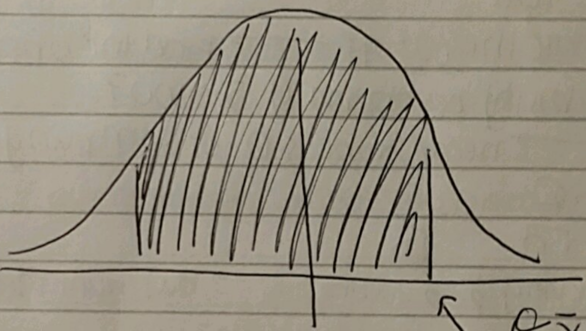
$H_0 : \mu = 77$
$H_1 : \mu \neq 77$

$$t = \frac{79.8 - 77}{1.34}$$

$\sigma_{\bar{x}}$ = SD of mean
(calculate SE)

$t = 2.09$
$df = N - 1$
$= 39 - 1$

* in terms of calculating the test statistic, everything will be the same



$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$

when $P(t) < \alpha = 0.05$, reject the null hypothesis
- Hypothesis is mutually exclusive
* can be either one-tailed or two-tailed
- 2 tailed test has less power

CONFIDENCE INTERVAL
- How can we tell if we reject the null hypothesis?
  $\bar{x} \pm t_{\alpha/2, df} * SE_{\bar{x}}$
- APA format → math?
  $t(29) = 2.09$
  $p = 0.023$

test → correcting mean
$MSE = \frac{var}{2}$

~~Two-sample test~~ ↗ find prob

- independent: randomly assigned

α = percentage of the p

TWO-SAMPLE T-TEST
- those two groups are equivalent — is it possible to observe this difference
- Null hypothesis

SE = average value?

- Homogeneity of variance → important assumption