Economics 2740
Department of Economics
University of Guelph

Tests About Population Means

Case 1: s^2 known, data normal

(Section 11.3 ed6 or 10.3 ed5)

- e.g. UofG isn’t cheap
- Is it worth it?
- Test if UofG grads earn more than other college grads, age 25
- Obviously not all UofG grads earn more – test if population mean is higher
- (Please note that the data for this example is fictitious)

How to test this

- Step 1: Look up average university grad salary age 25 in Statistics Canada ($\mu_0 = 50$ thousand CAD/yr.).
- Step 2: Survey random sample: "n" U of G grads age 25, and then simply record the salary levels of each respondent.

	Who	**Salary (thousands)**
X_1	Jason	55
X_2	Bob	45
X_3	Mary	60
X_4	Fumiko	60
X_5	Elvin	55
X_6	Anna	55

Step 3: **Calculate sample means**

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = 1/6(x_1 + x_2 + x_3 + x_4 + x_5 + x_6)$$

$$= 1/6(55 + 45 + 60 + 60 + 55 + 55)$$

$$= 55$$

Key Notes:

- 55 > 50
- This suggests UofG Grads do better
- However, this is based on just 6 people
- Did we get $\bar{X} > 50$ by chance?
- *Or* is it meaningful?

- To answer this need Hypothesis test

<u>Step 4:</u> **Specify H_0 and H_1**

<u>*Define*</u> μ = pop. Mean UofG Grad (age 25)

Now, let μ_0 = 50 (Hypothesized value)

Q: Can we reject $\mu \le 50$ *in favor of* $\mu > 50$*?*

H_0: $\mu \le \mu_0$ *(U of G Grads don't do better)*

H_1: $\mu > \mu_0$ *(U of G Grads do better)*

Step 5: **Set a *"Significance"* Level**

- What is the standard of proof
- How sure must we be before we say U of G Grads do better (i.e. reject H_0)

Follow Common Practice

- Must be 95% sure to reject
- Set α ***= 0.05***

  [i.e. Restrict Prob(Type I) $\leq$ 5%]

Step 6: **Form Test Statistic**

- *Intuition*: Reject if $\bar{X} - \mu_0 = \bar{X}$ - *50 large*

  - if $\bar{X}$ = 100 then $\bar{X}$ > 50 too big to be by chance

  - if $\bar{X}$ = 50.01 then $\bar{X}$ > 50 likely by chance

  - if $\bar{X}$ = 55 then hard to tell
- Q: How do we measure large?
- A: In "*Standard Deviations of* $\bar{X}$"
- Why: (1) Doesn't depend on the unit of measurement

  (2) Means same thing for all tests

- To do this, divide $\bar{X} - \mu_0$ by its standard deviation (SD)
- <u>Define:</u>

$$z = \frac{(\bar{x} - \mu_0)}{\sigma_{\bar{x}}} = \frac{(\bar{x} - \mu_0)}{\sigma / \sqrt{n}}$$

- Z measures the size of $\bar{X} - \mu_0$ in standard deviations
- If z = 3, then $\bar{X}$ is 3 standard deviations above 50
  - unlikely to happen by chance
  - suggests rejecting H_0

*W*hen H_0 holds and $\mu = \mu_0$, Then:

$$z == \frac{\left(\bar{x} - \mu_0\right)}{\sigma / \sqrt{n}} = \frac{\left(\bar{x} - \mu\right)}{\sigma / \sqrt{n}} \sim N(0,1)$$

Interpretation:

- If the hypothesized value is correct, then z centers the sample mean about true population mean
- In which case, z is standard normal

Step 7: Find the Critical Value

- Q: How many standard deviations above 50 must $\bar{X}$ be for us to reject?

  i.e. How large must z be?

- A: Z must be so large that the prob. of getting a z-value that big *when H_0 holds* is less than $a = 0.05$.

- The <u>critical value</u> Z_a is the cut-off point

  - if Z > Z_a then we Reject H_0

  - if Z ≤ Z_a then we Fail to reject H_0
- Choose Z_a so that:

  Prob (Type I Error) =

  Prob (Z > Z_a when $m = m_0$) = α
- <u>Define</u> Z_a by Prob (Z > Z_a) = α

  for Z ~ N(0,1)
- Please see graph of Normal distribution in the back of the book (i.e. $Z_{0.05} = 1.645$)

Step 8: Calculate Z and conduct test

- In our example, assume $\sigma = 2(6)^{1/2}$

$$z = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = \frac{55 - 50}{2\sqrt{6} / \sqrt{6}} = \frac{55 - 50}{2} = 2.5 > Z_a = 1.645$$

Reject H_0 at 5% significance level

It pays to go to Uof G! (At least in this sample)