Class 2: Hypothesis Testing for a Mean and Significance of Regression Coefficients

Table of Contents / Agenda

1 Review: Regression and Interpretation of Regression Results

$$y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \cdots + +\beta_K x_{i,K} \varepsilon_i$$

- Last class, we talked about how to interpret regression results

1.1 Model Level

- R^2: How much variation in y can my model explain?

$$R^2 = 1 - \frac{SS_{error}}{SS_{total}}$$

- We use **adjusted** R^2 to compare regressions with different numbers of independent variables

1.2 Variable Level (Coefficients)

- How does x_1 affects y?
- β_1: How y changes if x_1 is increased by 1 unit

1.3 Significance

- How do we know if those coefficients are *significant*?
- Does x_1 affects y?
  - We do hypothesis testing for each of the coefficients separately

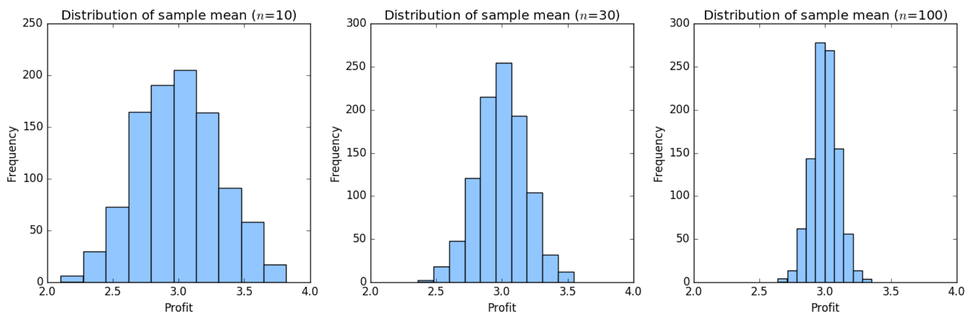
2 Sampling Distribution and Variance of the Sample Mean

- Suppose we can *repeatedly* draw random samples of size n from the population
- Then calculate the mean for *each* sample:
  - e.g., $\bar{x}_1$: mean of the first sample, $\bar{x}_2$: mean of the second sample, ...
- Computation of variance of these <u>means</u>:

$$s_{\bar{x}}^2 = \frac{s^2}{n} \begin{cases} \Uparrow & \text{increases with increasing sample variance}(s^2) \\ \Downarrow & \text{decreases with increasing sample size}(n) \end{cases}$$

2.1 How reliable is the sample mean as the sample size increases?

- The histogram of <u>sample means</u> will eventually look like a normal distribution when n becomes larger and larger. (Central-Limit Theorem)
- The histogram will also be "thinner" as n goes larger.



(these are generated with 1,000 draws of samples with size n - that is, we have 1,000 independent samples each with n observations)

2.2 Sample Mean and Standard Error

Sample mean $\overline{x}$

Accuracy of sample mean depends on:

Sample variance s_x^2

Sample size n

Standard Error $SE_{\overline{x}} = \sqrt{\frac{s_x^2}{n}}$

2.3 Standard Deviation VS. Standard Error

Standard deviation (s_x) Describes the spread of values in the sample

- The sample standard deviation, s_x is a random quantity - it varies from sample to sample
- Becomes **more accurate representation of the population standard deviation** when the sample size increases

Standard error of the mean ($SE_{\bar{x}}$) the standard deviation of the sample mean $\bar{x}$

- Describes $\bar{x}$'s accuracy as an estimate of the population mean, μ
- When the sample size increases, the estimator is based on more information and becomes more accurate, so **its standard error decreases**

2.4 Margin of error: Confidence interval of the mean

- What is a 95% confidence interval?
  - If we would draw samples of the same size over and over again, then in 95% of the times this interval covers the mean.
  - So it is a measure of reliability
  - $[\bar{x} - 1.96 SE_{\bar{x}}, \quad \bar{x} + 1.96 SE_{\bar{x}}]$
- It does not mean that "there is a 95% chance that the population mean is in this specific confidence interval" (See textbook page 277)

3 Hypothesis Testing for a Mean

Statements

- People think that the quality of food at our restaurant is above average (5)
- College head coaches' winning percentages affects their compensation levels (Regression)

3.1 Testing Hypothesis about a Single Mean

- People think that the quality of food at our restaurant is above average (5)
- Null hypothesis:
  - There is no difference or effect
    H_0: Average rating of food quality is 5 or, H_0: $Rating_{foodquality} = 5$
- Alternative hypothesis:
  - That there is a difference (or an effect)
  - Two-sided
    * The average rating of food quality is not 5
      H_a rating$_{foodquality} \neq 5$
    * No discrimination of the direction of difference
  - One-sided
    * The average rating of food quality is higher than 5
      H_a rating$_{foodquality} > 5$
    * Some idea of the direction of difference
- H_0: $rating_{foodquality} = 5$
- H_a: $rating_{foodquality} \neq 5$ or $rating_{foodquality} > 5$
- Average rating from 100 respondents: 6.5, Variance: 4
- Is the difference (5 VS. 6.5) statistically significant?

3.2 z-test vs. t-test

- Rule of thumb:

  "Use a z-test when the variance of the distribution is known, otherwise use a t-test"

- For a large sample, t-test is equivalent to a z-test
  - Mostly fine with Regression

As n increase, the t-distribution approaches the standard normal distribution

3.3 Recall

- Sample mean: $\overline{x}$
- Accuracy of sample mean depends on:
  - Sample variance: s_x^2

  - Sample size: n

- Standard Error: $SE_{\overline{x}} = \sqrt{\frac{s_x^2}{n}}$

Hence, with the example,

- $\overline{x} = 6.5, \qquad s_x^2 = 4, \qquad n = 100$
- $SE_{\overline{x}} = \sqrt{\frac{s_x^2}{n}} = \sqrt{\frac{4}{100}} = .2$

3.4 Computing t-statistic

- $H_0 : \overline{x} = \overline{x}_0 \qquad (6.5 = 5)$
- $H_a : \overline{x} \neq \overline{x}_0 \quad (6.5 \neq 5)$ or $H_a : \overline{x} > \overline{x}_0 \quad (6.5 > 5)$
- $t = \frac{\overline{x} - \overline{x}_0}{SE_{\overline{x}}}, \qquad$ where $\frac{\overline{x} - \overline{x}_0 = 1.5}{SE_{\overline{x}} = .2} \quad = 7.5$
- $t \uparrow \Rightarrow$ p-value $\downarrow$
  - p-value: the probability of finding the observed results when H_0 is true

3.5 t-critical

- How large must t be to reject the null hypothesis:
  - t must be larger than $t_{critical}$ (threshold)
  - Since the mean of t-distribution is zero, larger the value, smaller the probability
- $t_{critical}$ depends on:
  - Significance level α (typically $\alpha = 0.05$)
  - Degrees of freedom: total sample size-1 $\rightarrow n - 1$
  - Whether test is one sided or two sided
  - Use t-tables. For large n, you can use normal distribution (z)
- For two-sided test:
  - $t_{Critical} = t_{\alpha/2,n-1}$
    * For large n, $t_{0.025} = 1.96$
  - Reject the null if $|t| > t_{Critical}$
  - Fail to reject the null if $|t| < t_{Critical}$
- For one-sided test:
  - $t_{Critical} = t_{\alpha,n-1}$
    * For large n, $t_{0.05} = 1.65$
  - Reject the null if $t > t_{Critical}$
  - Fail to reject the null if $t < t_{Critical}$

- In both cases, rejecting the null means p-value $< \alpha$

3.6 The conclusion is

$$H_a : rating_{foodquality} \neq 5 \qquad\qquad H_a : rating_{foodquality} > 5$$

$$|t| = 7.5 > t_{critical} = 1.96 \Rightarrow \text{ Reject } H_0 \qquad\qquad t = 7.5 > t_{critical} = 1.65 \Rightarrow \text{ Reject } H_0$$

4 Significance of Regression Coefficients

- We do hypothesis testing for a mean for *each* coefficient separately

4.1 Null and Alternative Hypotheses

- The null hypothesis is that there is *no* effect:
  - $H_0 : \beta_k = 0 \qquad H_A : \beta_k \neq 0$
- Null hypothesis:
  - H_0: Winning Percentage of a head coach does not affect his compensation
  - H_0: $\beta_{WinPercentage} = 0$
- Alternative hypothesis:
  - College head coaches' winning percentages affects their compensation levels
  - H_a: $\beta_{WinPercentage} \neq 0$

4.2 Regression Results

- `t Stat`: $\frac{\hat{\beta}_k}{SE_k}$ (because $H_0 : \beta_k = 0$): Reject the null if $T > |1.96|$
- `P-value`: the probability of observing $\hat{\beta}_k$ if the null hypothesis is true: reject the null if `P-value` $< 0.05 (= \alpha)$
- 95% $((1 - \alpha) \times 100\%)$ Confidence interval: will not include 0 if $\hat{\beta}_k$ is significant