Tutorial: Hypothesis Testing

Hammad Shaikh

October 20, 2019

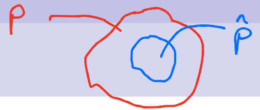
Hypothesis Testing

- A hypothesis is a statement about $\overbrace{\text{population parameters}}^{\mu,\ \sigma,\ p,\ \beta}$
  - Sample $\underbrace{\text{statistics}}_{\hat{\mu},\ \hat{\sigma},\ \hat{p},\ \hat{\beta}}$ don't belong in a hypothesis
- Null hypothesis: statement relating to the status quo
  - Example, H_0 : Teacher did not cheat

  Ex. $H_0: P_{tails} = \frac{1}{2}$ (fair coin)
- Alternative hypothesis: statement taking the opposite stance than the null hypothesis that the researcher is interested in
  - Example, H_1 : Teacher cheated

  Ex. $H_1: P_{tails} \neq \frac{1}{2}$ (Biased coin)

Hypothesis Test Example



- U of T has around 6000 students that enrol in ECO100. Suppose the dean claims that more than 80% of them complete the course. The dean asks you to test this claim. You have a survey of 500 students who initially enrolled in ECO100, 420 students report completing the course

(a) What is the population parameter of interest?

$$p = \frac{\#\text{ Complete ECO100}}{6000} \quad (\text{true prop. complete ECO100})$$

(b) How to estimate population parameter?

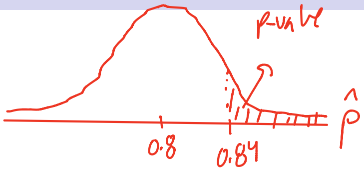
$$\hat{p} = \frac{\#\text{ Complete ECO100}}{500} \ (\text{Estimator}), \quad \tilde{p} = \frac{420}{500} = 0.84 \ (\text{Estimate})$$

(c) How to set up hypothesis test (H_0 and H_1)?

H_0: $p \leq 0.8$ $(p = 0.8)$

H_1: $p > 0.8$

Hypothesis Test Example Cont.



(d) What is the distribution of estimator?

Under H_0 $(P_0 = 0.8) \Rightarrow \hat{p} \sim N\left(P_0, \frac{P_0(1-P_0)}{n}\right)$

(e) What is p-value and conclusion?

p-value $= Pr(\hat{p} > \tilde{p}) = 0.009 < \alpha = 0.05$

$\alpha = 0.05$ $\Rightarrow$ rej. H_0

Conducting a Hypothesis Test

- Starting point is a $\overbrace{\text{estimator}}^{\hat{\mu},\,\hat{\sigma},\,\hat{p},\,\hat{\beta}}$ for the $\overbrace{\text{parameter(s) of interest}}^{\mu,\,\sigma,\,p,\,\beta}$
  - $\underbrace{\text{Realization}}_{\text{estimate}}$ from the estimator using a sample also required
- Assume H_0 is true
  - Identifies distribution for estimator $\widehat{X}$
- Compute the probability of obtaining a value for $\widehat{X}$ at least as extreme as that obtained from sample
  - This is known as the p-value
- Define significance level $\alpha \in \{0.01, 0.05, 0.1\}$
  - Fail to reject H_0 if p-value $> \alpha$
  - Reject H_0 if p-value $< \alpha$

Type I and Type II Errors

H_0: innocent
H_1: guilty

Conclude
sample

Truth	H_0	H_1
H_0	✓	Type II
H_1	Type I	✓

- We can never be 100% sure whether the conclusion obtained from the hypothesis test is correct
    - The conclusion may be incorrect (mistakes are possible)
- Type I Error: Rejecting a true null hypothesis (“false positive")
    - Hypothesis test says a honest teacher cheated
- Significance level $\alpha = \Pr(\text{Type I Error})$
- Type II Error: Failing to reject a false null (“false negative")
    - Hypothesis test says a cheating teacher did not cheat

$\downarrow$ Type I $\Rightarrow$ $\uparrow$ Type II (non-linear)

Simple Linear Regression and Hypothesis Testing

- Simple linear regression: $TestScore_d = \beta_0 + \beta_1 STR_d + \epsilon_d$
  - β_0 (intercept) and β_1 (slope) are unknown parameters
  - Use sample $(STR_d, TestScore_d)_{d=1}^n$ to make inference about the simple linear regression parameters

    $\hookrightarrow \hat{\beta}_1$

- Question: How much can we trust the primary estimate $\widehat{b}_1$?

- Null Hypothesis: Class size has no effects on achievement

  $\hookrightarrow H_0: \beta_1 = 0$

- Alternative Hypothesis: Class size effects achievement

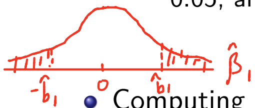
  $\hookrightarrow H_1: \beta_1 \neq 0$

SLR and Hypothesis Testing Cont.

$\hat{\beta}_1 = \frac{S_{xy}}{S_x^2}$

- Under $H_0 : \beta_1 = 0$ we have $\widehat{\beta}_1 \sim N(0, \frac{\sigma_\epsilon^2}{\sum_{i=1}^n (X_i - \bar{X})^2})$
  - Since ϵ unknown, σ_ϵ^2 is also unknown. The solution is to replace it with s_e^2, the sample variance of the residuals

- If $H_1 : \beta_1 \neq 0$ we compute p-value $= 2 * Pr(\widehat{\beta}_1 > \widehat{b}_1)$
  - $\widehat{b}_1$ is very significant if p-value < 0.01, significant if p-value < 0.05, and marginally significant if p-value < 0.1



- Computing p-value involves $SE(\widehat{b}_1) = \sqrt{Var(\widehat{\beta}_1)}$
  - Typically (not always) if $|\frac{\widehat{b}_1}{SE(\widehat{b}_1)}| > 2$ then $\widehat{b}_1$ is significant

$$t_{stat} = \frac{\hat{\beta}_1 - \beta_1}{SE(\hat{\beta}_1)}$$

Confidence Interval and Hypothesis Testing

$\mu = \text{pop}^n \text{ mean} \Rightarrow \begin{cases} \text{Point estimator: } \bar{X} \\ \text{Interval estimator: } \bar{X} \pm \underbrace{t_{\alpha/2, n-1} \frac{S_x}{\sqrt{n}}}_{\text{Margin Error}} \end{cases}$

- Confidence intervals: Estimate $\pm$ Margin of Error
  - Interval estimator for a population parameter
    - Probability interval estimator includes parameter is 1 - α

Ex. $Pr\left(\mu \in \left[\bar{X} - t_{\alpha/2, n-1} \frac{S_x}{\sqrt{n}}, \bar{X} + t_{\alpha/2, n-1}\right]\right) = 1 - \alpha$

- Fail to reject H_0 is CI estimate includes $\underbrace{\text{true parameter}}_{\text{Under } H_0}$

$H_0: \mu = 0 \Rightarrow 0 \in CI_\mu \Rightarrow FTR\ H_0$

- Reject H_0 is CI does not cover $\underbrace{\text{true parameter}}_{\text{Under } H_0}$

$H_0: \mu = 0 \Rightarrow 0 \notin CI_\mu \Rightarrow Rej.\ H_0$

Hypothesis Testing Example with Multiple Regression

Linear regression (wage = B0 + B1educ + B2motheduc + B3fatheduc + eps)

wage	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
educ	22.607	2.416	9.36	0.000	17.869	27.345	***
motheduc	5.237	2.335	2.24	0.025	0.659	9.816	**
fatheduc	3.636	1.959	1.86	0.064	-0.205	7.476	*
Constant	189.539	29.886	6.34	0.000	130.931	248.146	***

Mean dependent var	589.814	SD dependent var	265.115
R-squared	0.081	Number of obs	2220.000
F-test	65.053	Prob > F	0.000

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

- Is β_1 statistically significant at the 5% level?

$H_0: \beta_1 = 0$, $\hat{\beta}_1 = 22.6$, $\begin{cases} p \approx 0 < 0.05 \\ 0 \notin CI \end{cases} \Rightarrow$ rej. H_0

- Is β_3 statistically significant at the 5% level?

$H_0: \beta_3 = 0$, $\hat{\beta}_3 = 3.6$, $p = 0.064 > 0.05 \Rightarrow$ FTR H_0

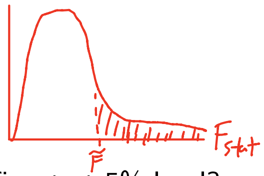
Testing for Overall Significance of Regression Model

- $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k$
    - $H_0 : \cancel{\beta_0} = \beta_1 = \ldots = \beta_k = 0$

    $H_1 : \beta_j \neq 0$ for $j \in \{1, \ldots, k\}$

    $k = \#$ of covariates

- $F_{stat} = \frac{R^2/k}{(1-R^2)/(n-k-1)} \sim F(k, n-k-1)$



- Is the regression on the previous slide significant at 5% level?

$p.\text{value} = Pr(F_{stat} > \tilde{F}) \approx 0 < \alpha \Rightarrow$ rej. H_0

Testing Significance For a Subset of Slope Parameters

R^2_{ur}

- *Unrestricted* : $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k$
  - $H_0 : \cancel{\beta_0} = \beta_1 = \ldots = \beta_q = 0$ for $q < k$

$H_1 : \beta_j \neq 0$ for $j \in \{0, 1, 2 \ldots, q\}$

R^2_r

- *Restricted* : $Y = \beta_{q+1} X_{\cancel{1}} + \beta_{\cancel{2}} X_{\cancel{2}} + \ldots + \beta_k X_k$

$q+1$ $q+2$ $q+2$

$q = \#$ of restrictions

- $F_{stat} = \frac{(R^2_{ur} - R^2_r)/q}{(1 - R^2_{ur})/(n-k-1)} \sim F(q, n-k-1)$