Development Statistics

S08 Statistical Test 1

Fujikawa, Kiyoshi
Nagoya University, GSID

Statistical Test

- To test a hypothesis
  - The academic ability of Japanese students are getting lower ?
  - Supporting ratio of ASO cabinet is decreasing ?
- But nobody knows the parameters of the population
  - Pick up samples
  - Estimate parameters based on samples

Sample & Population

- 標本と母集団 (人口ではない)
- Population parameters(母数)
  - Fixed number 定数
  - Population Mean 平均
  - Population variance 分散
- Samples (標本)
  - Random variable 確率変数
  - Sample mean 標本平均
  - Sample variance 標本分散

Low of large numbers

- 「大数の法則」
- If the sample size is large enough,
- 標本サイズが十分に大きければ
- The sample mean(variance) is asymptotically the population mean(variance) .
- 標本平均(分散)は漸近的に母集団の平均(分散)

Central limit theorem

- No matter what is the population is,
- 母集団がどんな分布であれ,
- Sample mean follows the Normal Distribution
- 標本平均は正規分布に従う

Test of population mean 1

- If σ^2 is known

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \Rightarrow z = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1)$$

- You can do a statistical test based on the standard normal distribution

σ^2 is known

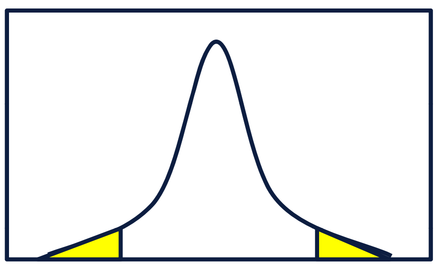
- Assumption on population & sample
  - $\sigma^2 = 100$
  - Sample size = 25
- Hypothesis
  - H0: population mean = 50
  - H1: population mean ≠ 50
- If H0 is true

$$z = \frac{\overline{X} - 50}{\sqrt{\frac{100}{25}}} = \frac{\overline{X} - 50}{2} \sim N(0,1)$$

σ^2 is known

- Significant level
  - The researcher's decision
  - 10%, 5%, 1%
- Suppose z=0.3
  - 0.3 is in a blue part
  - H0 is not rejected
- Suppose z=3.0
  - 3.0 is in a yellow part
  - H0 is rejected

Standard normal



-1.96 1.96

Test of population mean 2

- Even If σ^2 is unknown
- IF the sample size is large enough

$$z = \frac{\bar{X} - \mu}{\sqrt{\frac{s^2}{n}}} \sim N(0,1)$$

- You can do a statistical test based on the standard normal distribution

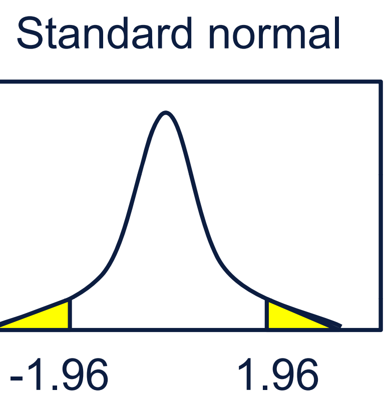
σ^2 is unknown, large sample

- Assumption on population & sample
  - $s^2 = 400$
  - Sample size = 100
- Hypothesis
  - H0: population mean = 50
  - H1: population mean ≠ 50
- If H0 is true

$$\frac{\bar{X} - 50}{\sqrt{\frac{400}{100}}} = \frac{\bar{X} - 50}{2} \sim N(0,1)$$

σ^2 is unknown, large sample

- Significant level
  - The researcher's decision
  - 10%, 5%, 1%
- Suppose z=0.3
  - 0.3 is in a blue part
  - H0 is not rejected
- Suppose z=3.0
  - 3.0 is in a yellow part
  - H0 is rejected



Test of population mean 3

- If σ^2 is unknown

$$t = \frac{\overline{X} - \mu}{\sqrt{\frac{s^2}{n}}} \sim t(n-1)$$

- You can do a statistical test based on student's t-distribution

σ^2 is unknown

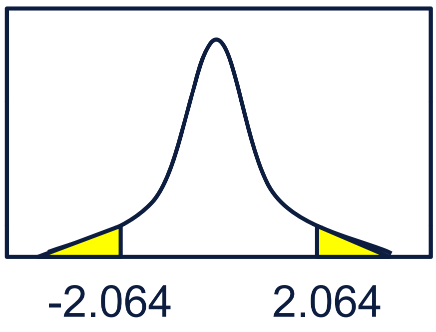
- Assumption on population & sample
  - s^2 =100
  - Sample size = 25
- Hypothesis
  - H0: population mean = 50
  - H1: population mean ≠ 50
- If H0 is true

$$\frac{\bar{X}-50}{\sqrt{\frac{100}{25}}}=\frac{\bar{X}-50}{2}\sim N(0,1)$$

σ^2 is known

- Significant level
  - The researcher's decision
  - 10%, 5%, 1%
- Suppose z=0.3
  - 0.3 is in a blue part
  - H0 is not rejected
- Suppose z=3.0
  - 3.0 is in a yellow part
  - H0 is rejected

T-dist(df:24)



Bernoulli distribution (Coin throwing)

- 1 (=success) with probability p (0.5)
- 0 (=failure) with probability 1-p (0.5)

- Mean

$$\mu = p*1 + (1-p)*0 = p = 0.5$$

- Variance

$$\sigma^2 = p*(1-p)^2 + (1-p)*(0-p)^2$$
$$= p(1-p) = 0.25$$

Test of population ratio

$$z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} \sim N(0,1)$$

- You can do a statistical test based on the standard normal distribution