Inferential Statistics, Test Statistics Manuel

Tingfeng Xia

2019, Winter Term

Contents

1 Test for $\mu = \mu_0$, w/σ^2 known

Assume that $X_i \sim N(\mu, \sigma^2)$ are i.i.d, then the test statistic is

$$T(X) = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

then under α significance level, we have the rejection region

$$R_\alpha(T) = (-\infty, z_{\frac{\alpha}{2}}) \cup (z_{1-\frac{\alpha}{2}}, \infty)$$

2 Test for $\mu = \mu_0$, w/σ^2 unknown

Assume that $X_i \sim N(\mu, \sigma^2)$ are i.i.d, then the test statistic is

$$T(X) = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$$

then under α significance level, we have the rejection region

$$R_\alpha(T) = (-\infty, t_{\frac{\alpha}{2}, df=n-1}) \cup (t_{1-\frac{\alpha}{2}, df=n-1}, \infty)$$

3 Test for $\sigma^2 = \sigma_0^2$

Assume that $X_i \sim N(\mu, \sigma^2)$ are i.i.d, then the test statistic is

$$T(X) = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{df=n-1}$$

and the α significance level rejection region is

$$R_\alpha(T) = (-\infty, \chi^2_{\frac{\alpha}{2}, df=n-1}) \cup (\chi^2_{1-\frac{\alpha}{2}, df=n-1}, \infty)$$

4 Equality of Variances $\sigma_x = \sigma_y$

If we have $X_1, \ldots, X_n \sim N(\mu_x, \sigma_x^2)$ and $Y_1, \ldots, Y_n \sim N(\mu_y, \sigma_y^2)$, then under our null hypothesis

$$T(X,Y) = \frac{S_x^2/\sigma_x^2}{S_y^2/\sigma_y^2} = \frac{S_x^2}{S_y^2} \sim F_{(n-1)(m-1)}$$

With α significance level, we then have the rejection region

$$R_\alpha(T) = \left(-\infty, F_{\frac{\alpha}{2}(n-1)(m-1)}\right) \cup \left(F_{1-\frac{\alpha}{2}(n-1)(m-1)}, \infty\right)$$

5 Equality of $\mu_x = \mu_y$, w/σ_x, σ_y known

If we have $\bar{X} \sim N\left(\mu_x, \frac{\sigma_x^2}{n}\right)$ and $\bar{Y} \sim N\left(\mu_y, \frac{\sigma_y^2}{n}\right)$, then

$$T(X,Y) = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}} \sim N(0,1)$$

6 Equality of $\mu_x = \mu_y$, w/$\sigma = \sigma_x = \sigma_y$ known

If this is the case, we can pull the $\sigma_x = \sigma_y = \sigma$ out from the above equation, we will have

$$T(X,Y) = \frac{\bar{X} - \bar{Y}}{\sigma\sqrt{(\frac{1}{n} + \frac{1}{m})}} \sim N(0,1)$$

7 Equality of $\mu_x = \mu_y$, w/$\sigma = \sigma_x = \sigma_y$ unknown

If we have $\bar{X} \sim N\left(\mu_x, \frac{\sigma_x^2}{n}\right)$ and $\bar{Y} \sim N\left(\mu_y, \frac{\sigma_y^2}{m}\right)$, then

$$T(X,Y) = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{(\frac{1}{n} + \frac{1}{m})}} \sim t_{n+m-2}$$

where, S_p is the polled sample variance, defined as

$$S_p^2 := \frac{(n-1)S_x^2 + (m-1)S_y^2}{n+m-2}$$

8 Equality of $\mu_x = \mu_y$, w/σ_x, σ_y unknown

In this case, we have a messy formula for the degrees of freedom, the test statistics that we use stays the same as above.

9 Equality of $\mu_x = \mu_y$ for paired data

We set the $H_0 : \mu_x - \mu_y = 0$, define $D = X - Y$, then $\mu_d = \mu_x - \mu_y$. Notice that $\mu_d = 0 \iff \mu_x = \mu_y$, then our test statistic is

$$T(D) = \frac{\bar{D}}{s_d/\sqrt{n}} \sim t_{n-1}$$

10 Restricted Likelihood Ratio Test

Define

$$\Lambda := \frac{max_{\theta \in \Omega_0}[L(\theta)]}{L(\hat{\theta})}$$

Denoting $p = \dim \Omega =$ number of free var in the whole space, and $d = \dim \Omega_0 =$ number of free var under our null hypothesis, we have

$$T(X) = -2 \ln \Lambda \xrightarrow{D} \chi^2_{df=p-d}$$

11 Unrestricted Likelihood Ratio Test for Equality of $\mu_x = \mu_y$ for Normally Distributed Random Variables

Consider i.i.d $X_1, \ldots, X_n \sim N(\mu_x, \sigma_x^2)$ and i.i.d $Y_1, \ldots, Y_m \sim N(\mu_y, \sigma_y^2)$. Notice that we have $p - d = 2 - 1 = 1$ in this case, and the likelihood is

$$L(\mu_x, \mu_y) = \left\{ \left(2\pi\sigma_x^2\right)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma_x^2} \sum_i (X_i - \mu_x)^2} \right\} \left\{ \left(2\pi\sigma_y^2\right)^{-\frac{m}{2}} e^{-\frac{1}{2\sigma_y^2} \sum_i (Y_i - \mu_y)^2} \right\}$$

by re-writing with $H_0 : \mu = \mu_x = \mu_y$, we have

$$L(\mu) = \left\{ \left(2\pi\sigma_x^2\right)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma_x^2} \sum_i (X_i - \mu)^2} \right\} \left\{ \left(2\pi\sigma_y^2\right)^{-\frac{m}{2}} e^{-\frac{1}{2\sigma_y^2} \sum_i (Y_i - \mu)^2} \right\}$$

Our test statistic is then

$$T(X, Y) = -2 \ln \Lambda = -2 \ln \frac{L(\hat{\mu})}{L(\hat{\mu_x}, \hat{\mu_y})} \sim \chi^2_{df=1}$$

where we remind ourselves that $\hat{\mu}_x = \bar{x}$ and $\hat{\mu}_y = \bar{y}$ here and $\hat{\mu}$ is some wright-ed average of $\bar{x}$ and $\bar{y}$ as below

$$\hat{\mu} = \left(\frac{\frac{1}{\sigma_x^2/n}}{\frac{1}{\sigma_x^2/n} + \frac{1}{\sigma_y^2/m}} \right) \bar{x} + \left(\frac{\frac{1}{\sigma_y^2/m}}{\frac{1}{\sigma_x^2/n} + \frac{1}{\sigma_y^2/m}} \right) \bar{y}$$

12 Chi-Square Test of Goodness of Fit

Suppose, $X_1, X_2, \ldots, X_k$ are the observed counts of category $1, 2, \ldots, k$ respectively. Then

$$(X_1, X_2, \ldots, X_k) \sim \text{Mult}(n, p_1, p_2, \ldots, p_k) \quad \text{where } E[X_i] = np_i (\geq 1), \forall i$$

and our test statistic will be, in this case

$$T(X) = X^2 = \sum_{i=1}^{k} \frac{(X_i - np_i)^2}{np_i} \xrightarrow{D} \chi^2_{(df=k-1)}$$

13 Discrepancy Statistic for Normal R.V.s

Consider $X_1, \ldots, X_n \sim N(\hat{\mu}, \sigma_0^2)$ where σ_0^2 is known. Define $R = X_i - \bar{X}$, where $R \sim N(0, \sigma_0^2(1 - \frac{1}{n}))$ then, the descrepancy statistic is defined as

$$D(R) = \frac{1}{\sigma_0^2} \sum_{i=1}^{n} R_i^2 = \frac{1}{\sigma_0^2} \sum_{i=1}^{n} (X_i - \bar{X})^2 \sim \chi_{n-1}^2$$