VE401 Probabilistic Methods in Eng.

RC#5 Basic Test Theory, Tests on Normal Distribution

> By TA: DONG Juechu, Mar. 2021

if you want to edit this note, you can find it here https://github.com/joydddd/VE401-2020SP-notes

Tests

Hint

No need to carefully exam the logistic of the tests (Many of them are not logical, as Horst may say, "bullshit tests")

Start to use Mathematica! Write tests in a notebook so that you can use them in your final exam. You're expected to be familiar with it.

> **To get Access***
>
> (i) Visit https://user.wolfram.com/portal/registration.html and create a Wolfram ID. You must use an @sjtu.edu.cn email address and give your first and last names in pinyin (example: Xu Baishen enters last name: Xu and first name: Baishen).
> (ii) Next, visit
> https://user.wolfram.com/portal/requestAK/c51e79e5334a3600a4f740a2b3720961216dbc17
> and request an Activation Number. Make a note of the activation number. You will be directed to a page where you can download the installation binaries for the most current version of Mathematica. (You must select whether you want those for Windows, Linux or OS X.) The software binaries are several GB in size; it may perhaps be possible to share them amongst yourselves to
> save download time. Try it out and let me know.
>
> (iii) After downloading, you can install the software. You will be asked to enter the Activation Number you noted above and you will need internet access. Mathematica will then run on a temporary two-week license. Your name will be checked against a list, and if successful, the license will automatically be extended for one year. Therefore, it is very important that you enter your name properly when you request the Wolfram ID.
>
> *from VV286 FA19 slides by Horst Hohberger

Review the example on slides carefully -- as long as you know how to follow them step by step, you know how to pass the final exam

Goal

To reject or prove a hypothesis at the confidence level of P.

Fisher's Test

Denote the hypothesis by H_0.

One tail test

$H_0 : \theta \leq \theta_0 \quad \text{or} \quad H_0 : \theta \geq \theta_0$

P-Value = 0.95 if not explicitly stated. D is test statics.

if ($P[D|H_0] \leq P - value]$)

we reject H_0 at the P-value level of significance

else we fail to reject H_0

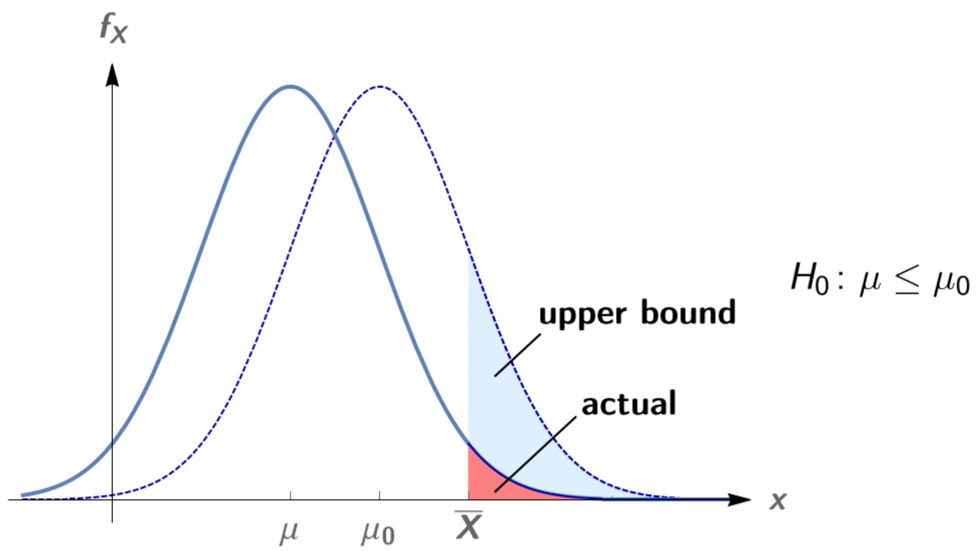
aka: We don't know if H_0 is true. We make no conclusion.

Based on Normal distribution

Sample mean is normally distributed with expectation μ and standard deviation $\sigma/\sqrt{n}$

-> calculate $\alpha(pValue) = P[Z < z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}]$ try to do it with mma!

If $\mu < \mu_0$, the probability will be smaller, since the density curve will be shifted to the left.



P[]

Two-Tailed Test

$H_0 : \theta \leq \theta_0$

We then find the probability of obtaining the measured value of $\bar{x}$ or a larger result if $\theta = \theta_0$. This is said to be the significance or P -value of the test.

-> For normal distribution, $\alpha(pValue)/2 = P[Z < z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}]$

Neyman-Pearson Decision Theory

A forced decision

Error Model

- We reject H0 (and accept H1) when H1 is true.
- We reject H0 (accept H1) even though H0 is true **(Type I error)**.
  - $\alpha := P[\,\text{Type I} \mid \text{error}\,] = P[\,\text{reject } H_0 \mid H_0 \text{ true}\,] = P[\,\text{accept } H_1 \mid H_0 \text{ true}\,]\,.$ **(confidence)**
- We fail to reject H0 even though H1 is true **(Type II error)**.
  - $\beta := P[\,\text{Type II} \mid \text{error}\,] = P[\,\text{fail to reject } H_0 \mid H_1 \text{ true}\,] = P[\,\text{accept } H_0 \mid H_1 \text{ true}\,]\,.$ **(power)**
- We fail to reject H0 when H0 is true.

Theoretically, H1 and ~H0 are different. In practice ... often H1 = ~H0. In this case, α is the same as P-Value in Fisher's test.

Critical Region and α

For certain α value, decide the region we take H0 if the static falls in it.

> in practice: reverse calculation of Fisher's Test
>
> Fisher: know statics calculate α
>
> Critical region: for given α calculate upper(or and lower) limit of statics

Normal distribution

- (two tailed)
  - Possibility of $\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}} > z_{\alpha/2}$ is equal to α. Therefore, the critical region is determined by

$$\bar{x} \neq \mu_0 \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Power and β

$$H_0 : \mu = \mu_0, \quad H_1 : |\mu - \mu_0| \geq \delta_0$$

$$-z_{\alpha/2} \leq Z \leq z_{\alpha/2}$$

$$\mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \bar{x} < \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$n \to \beta$

$$\beta = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-z_\beta} e^{-t^2/2}dt$$

$\beta \to n$ (sample size)

$$-z_\beta \approx z_{\alpha/2} - \delta\sqrt{n}/\sigma$$

or

$$n \approx \frac{\left(z_{\alpha/2} + z_{\beta}\right)^2 \sigma^2}{\delta^2}$$

(Use OC curve or mma, don't use this estimation unless this the only way to do it)

> Exercise 5.4 NHST Supercavitation is a propulsion technology for undersea vehicles that can greatly increase their speed. It occurs above approximately 50 meters per second, when pressure drops sufficiently to allow the water to dissociate into water vapor, forming a gas bubble behind the vehicle. When the gas bubble completely encloses the vehicle, supercavitation is said to occur.
>
> Eight tests were conducted on a scale model of an undersea vehicle in a towing basin with the average observed speed $\bar{x} = 102.2$ meters per second. Assume that speed is normally distributed with known standard deviation $\sigma = 4$ meters per second.
> i) Test the hypothesis $H_0 : \mu \leq 100$ versus $H_1 : \mu > 100$ using $\alpha = 0.05$. (2 Marks)
> ii) What is the P -value for the test in part (i)? (1 Mark)
> iii) Find the power of the test if the true mean speed is as low as 105 meters per second. (1 Mark)
> iv) What sample size would be required to detect a true mean speed as low as 105 meters per second if we wanted the power of the test to be at least 0.85? (1 Mark)
> v) Explain how the question in part could be answered by constructing a one-sided confidence bound on the mean speed. (1 Mark)

Tests on Normal distribution

T-Test

Normal distribution with **unknown variance**, test on **mean.**

$$T_{n-1} = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

We reject at significance level α

$$\begin{aligned}
&\text{- } H_0 : \mu = \mu_0 \text{ if } |T_{n-1}| > t_{\alpha/2,n-1}, \\
&\text{- } H_0 : \mu \leq \mu_0 \text{ if } T_{n-1} > t_{\alpha,n-1} \\
&\text{- } H_0 : \mu \geq \mu_0 \text{ if } T_{n-1} < -t_{\alpha,n-1}.
\end{aligned}$$

OC Curve

$$d = \frac{|\mu - \mu_0|}{\sigma}, \quad \text{approximation: } \sigma = S$$

Chi-Square Test

Normal distribution with unknown variance, test on **variance**

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

We reject at significance level α

- $H_0 : \sigma = \sigma_0$ if $\chi^2_{n-1} > \chi^2_{\alpha/2,n-1}$ or $\chi^2_{n-1} < \chi^2_{1-\alpha/2,n-1}$
- $H_0 : \sigma \leq \sigma_0$ if $\chi^2_{n-1} > \chi^2_{\alpha,n-1}$
- $H_0 : \sigma \geq \sigma_0$ if $\chi^2_{n-1} < \chi^2_{1-\alpha,n-1}$

Non-parametric Statistics

doesn't depend on any parametric and is **distribution-free**

gives estimation of **median or quartile** instead of mean or variance

Sign Test

Test for median

$$Q_+ = \#\{X_k : X_k - M_0 > 0\}, \quad Q_- = \#\{X_k : X_k - M_0 < 0\}$$

We reject at significance level α

- $H_0 : M \leq M_0$ if $P[Q_- \leq k \mid M = M_0] < \alpha$
- $H_0 : M \geq M_0$ if $P[Q_+ \leq k \mid M = M_0] < \alpha$
- $H_0 : M = M_0$ if $P[\min(Q_-, Q_+) \leq k \mid M = M_0] < \alpha/2$

Wilcoxon Signed Rank Test

Assumption: The distribution is symmetric about its median

rank: rank of absolute difference between data and hypothesis

Tie resolution: rank = average of their ranks.

$$W_+ = \sum_{R_i > 0} R_i, \quad |W_-| = \sum_{R_i < 0} |R_i|$$

Ne reject at significance level α

- $H_0 : M \leq M_0$ if W_- is smaller than the critical value for α,
- $H_0 : M \geq M_0$ if W_+ is smaller than the critical value for α
- $H_0 : M = M_0$ if $W = \min(W_+, |W_-|)$ is smaller than the critical value for $\alpha/2$

Find critical value in table unless you are explicitly told to use a normal approximation.

Normal Approximation

when $n \geq 10$ and assume it follows normal distribution, we can approximate critical value using normal distribution of

$$E[W] = \frac{n(n+1)}{4}, \quad Var[W] = \frac{n(n+1)(2n+1)}{24} - tiecorrection$$

for each group of t ties, the variance is reduce by $(t^3 - t)/48$

> Note: this approximation is quite rough and in the case of many ties, the variance can even be less than 0!