Hypothesis Testing - Handwritten Notes | MATH 312

Probability and Statistics

Drexel University

20 pag.

WEEK 5

Note Title 2/13/2007

Hypothesis testing

Terminologies:

NULL HYPOTHESIS (usually denoted by H_0)
— "prior belief"

ALTERNATIVE HYPOTHESIS (denoted by H_a)
— "alternative belief"

typical situation : Collect data, use the data to decide whether H_0 should be rejected.

1. A TEST STATISTIC is a function of the sample data on which the decision is to be based.

2. A REJECTION REGION is the set of all test statistic values for which H_0 will be rejected.

A TYPE I ERROR – rejecting the null hypothesis H_0 when it is true

A TYPE II ERROR – not rejecting H_0 when H_0 is false.

E.g.

Two companies

p = proportion of all potential subscribers who favor the 1st company over the 2nd

$H_0 : \quad p = 0.5$

$H_a : \quad p \neq 0.5$

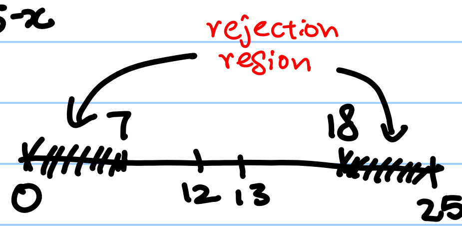
X = # of individuals in a random sample of size 25 individuals who favor the 1st company

x - an observed value of X

(a) Rejection region $R = \{x : x \leq 7 \text{ or } x \geq 18\}$

(c) Distribution of X when H_a is true (i.e. $p = 0.5$)

$$P[X=x] = \binom{25}{x}\left(\tfrac{1}{2}\right)^x\left(1-\tfrac{1}{2}\right)^{25-x}$$



$\alpha = P[\text{Type I error}] = ?$

$= P[X \leq 7 \text{ or } X \geq 18]$

$= \sum_{x=0}^{7} b(x; 25, \tfrac{1}{2}) + \left(1 - \sum_{x=0}^{17} b(x; 25, \tfrac{1}{2})\right)$

Table A.1 pg 738 $\quad 0.022 + 1 - 0.978$

$= 0.044$

d. $\beta(p) = P[\text{Type II error when the true } p \text{ is } p \neq 0.5]$

Note: Distribution of X when $p=0.3$ is

$$P[X=x] = \binom{25}{x}(0.3)^{x}(1-0.3)^{25-x}$$

$$\begin{aligned}\beta(0.3) &= P[\text{ we fail to reject } H_0 \text{ when } p=0.3] \\ &= P[8 \leq X \leq 17] = P[X \leq 17] - P[X \leq 5] \\ &= \sum_{x=0}^{17} b(x;25,0.3) - \sum_{x=0}^{7} b(x;25,0.3)\end{aligned}$$

$\underset{\text{Table A.1}}{=}$ ________ $-$ ________ $=$ ____

Tests about a population mean (Section 8.2)

3 cases (as in the study of CI)

(I) normal population with known σ

(II) not neccessarily normal, but n large

(III) normal population, σ not known, n not necessarily large

(I) $N(\mu, \sigma)$ – σ known

want to decide if we should reject the null hypothesis $H_0 : \mu = \mu_0$ or not.

we have a random sample

$$X_1, X_2, \cdots, X_n \sim_{iid.} N(? = \mu, \sigma)$$

Recall: if μ is really μ_0, then

$$\bar{X} = (X_1 + \cdots + X_n)/n \sim N(\mu_0, \sigma/\sqrt{n})$$

or

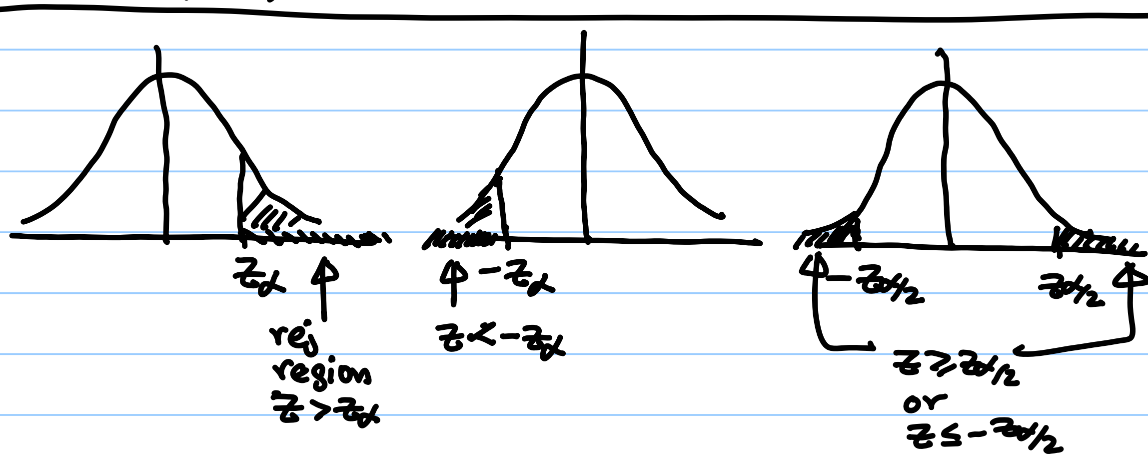
$$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

One sample z test

NULL Hypothesis H_0 : $\mu = \mu_0$

Test statistic value : $z = \bar{x} - \mu_0 / \sigma/\sqrt{n}$

Alt. hypothesis	Rejection region for level α test
$H_a : \mu > \mu_0$	$z \geq z_\alpha$
$H_a : \mu < \mu_0$	$z \leq -z_\alpha$
$H_a : \mu \neq \mu_0$	$z \geq z_{\alpha/2}$ or $z \leq -z_{\alpha/2}$



Q: What's the rationale behind these?

A: The rejection regions are chosen in such a way that

$$P[\text{Type I error}] = \alpha$$

Note : $\alpha \downarrow \Rightarrow$ rejection region smaller

(makes sense, right?)

What about Type II error?

say, if μ is not μ_0 but is μ' $(>\mu_0)$
(so we should reject H_0 - assume $H_a: \mu>\mu_0$)

What is P[type II error]?

accept H_0 when $\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < z_\alpha$

$$= P[\mu=\mu'\neq\mu_0 \text{ but the Level } \alpha \text{ test } \underbrace{\boxed{\text{fails to reject}}}_{=\text{ accepts}} H_0]$$

$$= P\left[\bar{X} < \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}} \text{ when } \mu=\mu'\right]$$

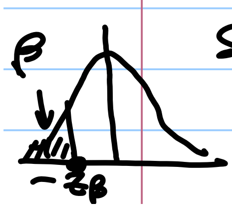
Note: $\mu=\mu' \Rightarrow \bar{X} \sim N(\mu', \sigma/\sqrt{n})$

$$= P\left[\underbrace{\boxed{\frac{\bar{X}-\mu'}{\sigma/\sqrt{n}}}}_{\sim N(0,1)} < \frac{\mu_0-\mu'}{\sigma/\sqrt{n}} + z_\alpha \text{ when } \mu=\mu'\right]$$

$$= \Phi\left(z_\alpha + \frac{\mu_0-\mu'}{\sigma/\sqrt{n}}\right)$$

Note:
- $\mu' \uparrow \Rightarrow$ P[type II error] $\downarrow$
- $\sigma \uparrow \Rightarrow$ P[type II error] $\uparrow$
- $n \uparrow \Rightarrow$ P[type II error] $\downarrow$

How large should n be such that P[type II error] $=\beta$?



Solve: $\Phi\left(z_\alpha + \dfrac{\mu_0-\mu'}{\sigma/\sqrt{n}}\right) = \beta$

or $-z_\beta = z_\alpha + \dfrac{\mu_0-\mu'}{\sigma/\sqrt{n}}$ or $n = \left[\dfrac{\sigma(z_\alpha+z_\beta)}{\mu_0-\mu'}\right]^2$

$\boxed{\Phi \text{ standard normal cdf}}$

Alternative hypothesis	Type II error prob. $\beta(\mu')$ for a Level α test
$H_a: \mu > \mu_0$	$\Phi\left(z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$
$H_a: \mu < \mu_0$	$1 - \Phi\left(-z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$
$H_a: \mu \neq \mu_0$	$\Phi\left(z_{\frac{\alpha}{2}} + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right) - \Phi\left(-z_{\frac{\alpha}{2}} + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$

Sample size n for which a level α test also has $\beta(\mu') = \beta$ at the alternative value μ' is

$$n = \begin{cases} \left[\dfrac{\sigma(z_\alpha + z_\beta)}{\mu_0 - \mu'}\right]^2 & \left(H_a: \mu > \mu_0 \text{ or } H_a: \mu < \mu_0\right) \\ \left[\dfrac{\sigma(z_{\frac{\alpha}{2}} + z_\beta)}{\mu_0 - \mu'}\right]^2 & \left(H_a: \mu \neq \mu_0\right) \end{cases}$$

Go through example 8.6 and 8.7 in class

(II) n large (underlying population distribution needs not be normal)

$$\boxed{\text{Recall } n \geq 40, \quad \frac{\bar{X}-\mu}{S/\sqrt{n}} \text{ is approximately } N(0,1)}$$

So can use the tests in case (I) with the appropriate changes
(basically just change σ to $S=\sqrt{\frac{1}{n-1}\sum_{1}^{n}(x_i-\bar{x})^2}$.

Read example 8.8

(III) n not necc. large, normal population

Recall $T = \dfrac{\bar{X}-\mu}{S/\sqrt{n}}$ has a t-distribution with $(n-1)$ d.o.f

The one-sample t-test

Null hypothesis : $H_0 : \mu = \mu_0$

Test statistic value : $t = (\bar{x}-\mu_0)/(s/\sqrt{n})$

Alternative hypothesis	Rejection region for a level α test
$H_a : \mu > \mu_0$	$t \geq t_{\alpha, n-1}$
$H_a : \mu < \mu_0$	$t \leq -t_{\alpha, n-1}$
$H_a : \mu \neq \mu_0$	$t \geq t_{\frac{\alpha}{2}, n-1}$ or $-t_{\frac{\alpha}{2}, n-1}$

Same reasoning behind the z-test :

e.g.

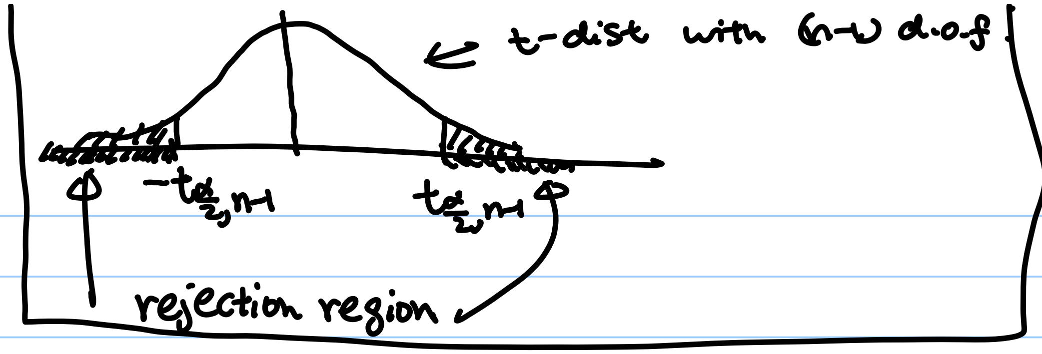
$$P[\text{type I error}] \overset{?}{=} \alpha \leftarrow \text{the "}\boxed{\text{tolerance}}\text{" for type I error}$$

assume $H_a : \mu \neq \mu_0$

$$\| $$

$$P[T \in \text{rejection region when } \mu = \mu_0]$$

So if we choose the rejection region to be ;



then we get what we want!

Determing β in this case is more complicated. (can you see why?)

see P 302

Read Example 8.9

Tests concerning a population proportion (Sec 8.3)

e.g.

$p = P[\text{"G"}]$

Hypothesis testing:

$H_0: p = 0.4 = p_0 \leftarrow$ null value

$H_a: p > 0.4$

etc.

1 if 'G'
0 if 'B'

iid. Bernoulli r.v.'s $X_1, \ldots, X_n$ with parameter p

$\hat{p} = (X_1 + \cdots + X_n)/n$ $\qquad E[\hat{p}] = p$, $\quad VAR[\hat{p}] = \dfrac{p(1-p)}{n}$

Test statistic:

$$\frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} \quad \text{(approximately } N(0,1)\text{)}$$

Null hypothesis : $H_0: p = p_0$

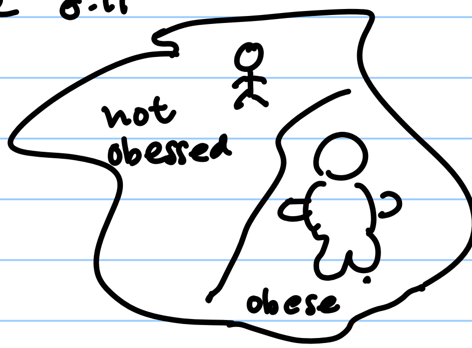
test statistic value : $z = (\hat{p} - p_0)/\sqrt{p_0(1-p_0)/n}$

Alt. hypothesis	rejection region (level α)
$H_a: p > p_0$	$z \geq z_\alpha$
$H_a: p < p_0$	$z \leq -z_\alpha$
$H_a: p \neq p_0$	$z \geq z_{\alpha/2}$ or $z \leq -z_{\alpha/2}$

(Note : test relies on CLT, use only when $np_0 \geq 10$ and $n(1-p_0) \geq 10$)

$$\boxed{P[\text{type I error}] \approx \alpha}$$

Example 8.11



sample
4115 adults
↳ 1276 obese

$H_0: p = 0.3$

$H_a: p > 0.3$

Hyp. testing with $\alpha = 0.1$

$4115(0.3) > 10$, $4115(0.7) > 10$ (large sample test can be used)

test statistic value

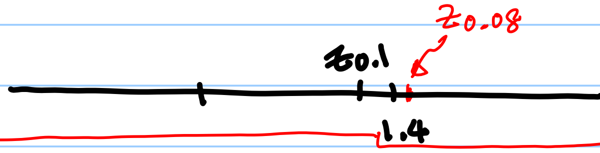
$$z = (\hat{p} - 0.3)(\sqrt{(0.3)(0.7)/n})$$

$$= \left(\frac{1276}{4115} - 0.3\right) \Big/ \sqrt{(0.3)(0.7)/4115}$$

$$= 1.40.$$

$z_{0.1} = 1.28$

since $1.40 > z_{0.1}$, the test suggests to reject H_0



Q: What if $\alpha = 0.08$, $z_{0.08} = \Phi^{-1}(0.92) = 1.4\ldots > 1.4$

Since $1.40 < z_{0.08}$, the test does not reject H_0

β and sample size determination

What if H_0 is not true, i.e.

$$P = p' \neq P_0$$

test statistic

$$Z = \frac{\hat{p} - P_0}{\sqrt{P_0(1-P_0)/n}} \text{ is no longer } N(0,1)$$

Just like before, in order to calculate $P[\text{type II error}]$, we need to first understand the distribution of Z when the true p is p'.

$$\hat{p} = \frac{X_1 + \cdots + X_n}{n} \qquad X_i \overset{iid}{\sim} \text{Bernoulli}(p')$$

n large

so $\hat{p}$ is approximately $N\left(p', \frac{\sqrt{p'(1-p')}}{\sqrt{n}}\right)$

and Z is approximately $N\left(\frac{p' - P_0}{\sqrt{P_0(1-P_0)/n}}, \sqrt{\frac{p'(1-p')}{P_0(1-P_0)}}\right)$

(Exercise : why?)

Alt. Hypothesis	$\beta(p')$
$H_a: p > p_a$	$\Phi\left(\dfrac{p_0 - p' + z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right)$
$H_a: p < p_a$	$1 - \Phi\left(\dfrac{p_0 - p' - z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right)$
$H_a: p \neq p_a$	$\Phi\left[\dfrac{p_0 - p' + z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right] - \Phi\left[\dfrac{p_0 - p' - z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right]$

(Exercise: derive these yourself)

If α, β are prespecified, minimum sample size to achieve these:

$$n = \begin{cases} \left[\dfrac{z_\alpha\sqrt{p_0(1-p_0)} + z_\beta\sqrt{p'(1-p')}}{p' - p_0}\right]^2 & \text{one tailed test} \\ \left[\dfrac{z_{\alpha/2}\sqrt{p_0(1-p_0)} + z_\beta\sqrt{p'(1-p')}}{p' - p_0}\right]^2 \end{cases}$$

Example 8.12 (sample exam question !)

A package-delivery company advertises

" at least 90% of packages are delivered on time "

p = "true proportion of packages delivered on time"

$p = 0.8$ [only the company would know this secret, of course.]

Q: $\boxed{\text{Based on a level 0.01 test with } n = 225 \text{ packages, How likely can the test successful tell that the company is lying in the ad. ?}}$

$H_0: p = 0.9$ (the company's claim)

$H_a: p < 0.9$

$$\beta(0.8) = P[\text{level 0.01 test } \boxed{\text{fails}} \text{ to reject } H_0 \text{ when } p = 0.8]$$

$$= 1 - \Phi\left[\frac{0.9 - 0.8 - \overset{2.33}{\overset{\shortparallel}{z_{0.01}}} \sqrt{(0.9)(0.1)/225}}{\sqrt{(0.8)(0.2)/225}}\right]$$

$$= 1 - \Phi(2.00) = 0.0228$$

So $P[\text{level 0.01 test with } n = 225 \text{ can reject } H_0] = 1 - \beta(0.8)$
$= 0.9772.$

If we want $\alpha = \beta = 0.01$

$$z_\alpha = z_\beta = 2.33$$

$$n = \left[\frac{2.33\sqrt{(0.9)(0.1)} + 2.33\sqrt{(0.8)(0.2)}}{0.8 - 0.9} \right]^2$$

$$\approx 266.$$

Small sample test for p :

use binomial distribution

(instead of a CLT approx.)

see pg 309

.

P- values (section 8.4)

Consider the following scenario:

- 10 = true average time to initial relief of of pain of a best-selling pain reliever
- the company considers introducing a newly developed reliever

  μ = average time for the new pain reliever

$H_0 : \mu = 10$ $H_a : \mu < 10$

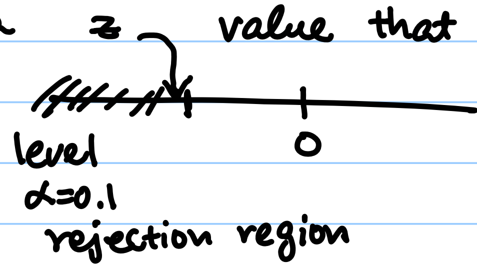
Company tells customers

"We reject H_0 (i.e. new reliever is better) based on a hypothesis testing analysis with $\alpha = 0.10$."

A customer satisfied with the best-selling reliever would view a Type I error serious. He/she (contemplating a switch to the new reliever) would question the validity of the company's claim; in particular, he/she may want to test the hypothesis with $\alpha = 0.05, 0.01$, or even lower.

Problem: The company's statement prevents an individual decision maker from reaching a conclusion at such a level (of α).

Related problem: may be the data gathered by the company gives a z value that barely falls into the rejection region.



H_0 is rejected (using $\alpha = 0.1$), but very marginally, and the customers do not know it from the statement (*).

Can the company say something to the public which is more informative?

There is a nice solution to this, all you need is to understand a simple fact :

[Read Ex. 8.5 and 8.15]

Example

$H_0 : \mu = 1.5$

$H_a : \mu > 1.5$

(σ known in this case) $= 0.20$

$x_1, \ldots, x_{32}$ given

$$\bar{x} = \frac{x_1 + \cdots + x_{32}}{32} = 1.5742$$

$$\text{test statistic value} = \frac{\bar{x} - 1.5}{(0.2)/\sqrt{32}} = 2.10$$

more and more "picky" about type I error

Level of significance α	rejection region	Conclusion
0.05	$z \geq 1.645$	reject H_0
0.025	$z \geq 1.96$	reject H_0
0.01	$z \geq 2.33$	do not reject H_0
0.005	$z \geq 2.58$	do not reject H_0

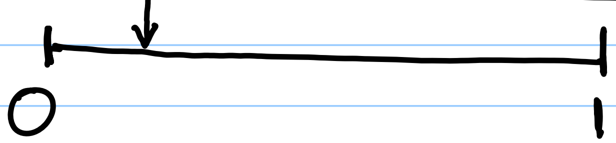
KEY OBSERVATION: There is a critical α value α^* at which we switch from reject to not reject.

called "P-value"

$\alpha \geq$ P-value $\Rightarrow$ reject H_0 at level α

$\alpha <$ P-value $\Rightarrow$ do not reject H_0 at level α

$\boxed{\text{Pvalue} = \text{smallest level at which } H_0 \text{ can be rejected}}$



For a z-test (z = test statistic value)

$$P = \begin{cases} 1 - \Phi(z) & \text{for an upper-tailed test} \\ \Phi(z) & \text{for a lower-tailed test} \\ 2(1 - \Phi(|z|)) & \text{for a two-tailed test} \end{cases}$$

Read Ex 8.17

For a t-test,

cdf of $N(0,1)$

Change the $\Phi(\ )$ above to the c.d.f. of the t-distribution with $(n-1)$ d.o.f. (n = sample size.)

Read Ex 8.18