Math 362: Mathematical Statistics II

Le Chen
le.chen@emory.edu

Emory University
Atlanta, GA

Last updated on April 13, 2021

2021 Spring

Chapter 6. Hypothesis Testing

Chapter 6. Hypothesis Testing

	True State of Nature	
	H_0 is true	H_1 is true
Fail to reject H_0	Correct	Type II error
Reject H_0	Type I error	Correct

<table>
<tr><th colspan="2" rowspan="2">Table of error types</th><th colspan="2">Null hypothesis (H_0) is</th></tr>
<tr><th>True</th><th>False</th></tr>
<tr><td rowspan="2">Decision about null hypothesis (H_0)</td><td>Don't reject</td><td>Correct inference
(true negative)
(probability = 1 - α)</td><td>Type II error
(false negative)
(probability = β)</td></tr>
<tr><td>Reject</td><td>Type I error
(false positive)
(probability = α)</td><td>Correct inference
(true positive)
(probability = 1 - β)</td></tr>
</table>

Type I error $\sim \alpha$

$$\alpha := \mathbb{P}(\text{Type I error}) = \mathbb{P}(\text{Reject } H_0 | H_0 \text{ is true})$$

By convention, H_0 is always of the form, e.g., $\mu = \mu_0$. So this probability can be exactly determined. It is equal to the level of significance α.

(Simple null test)

Type II error $\sim \beta$

$$\beta := \mathbb{P}(\text{Type II error}) = \mathbb{P}(\text{Fail to reject } H_0 | H_1 \text{ is true})$$

In order to compute Type II error, we need to specify a concrete alternative hypothesis.

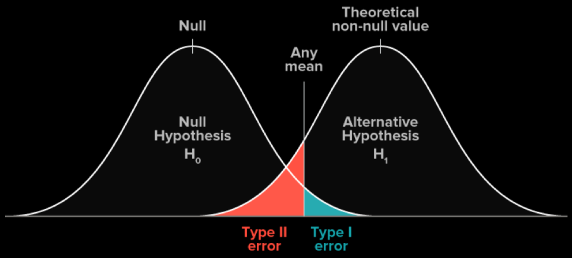


Figure: One-sided inference $H_1 : \mu > \mu_0$

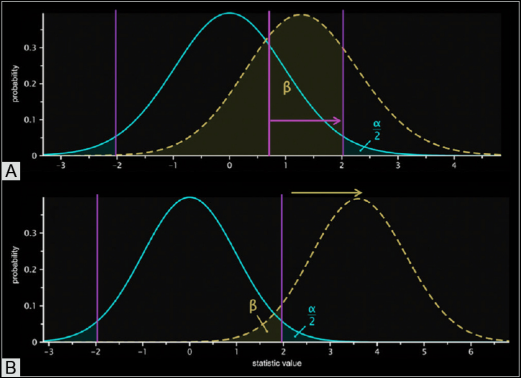
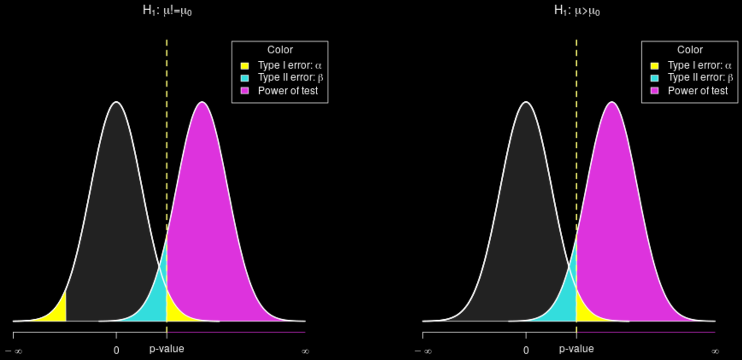


Figure: Two-sided inference $H_1 : \mu \neq \mu_0$

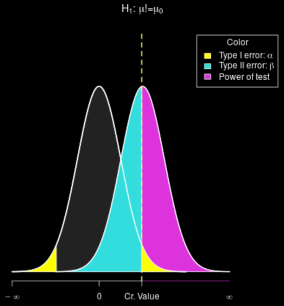
Power of test $1-\beta$

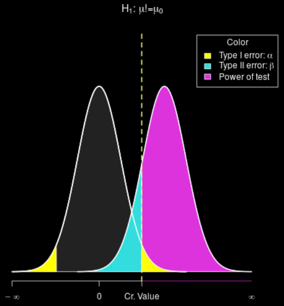
$$\text{Power of test} = \mathbb{P}(\text{Reject } H_0 | H_1 \text{ is true}) = 1-\beta$$

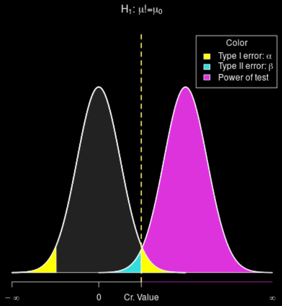


One online interactive show all α, β and $1-\beta$:
`https://rpsychologist.com/d3/NHST/`

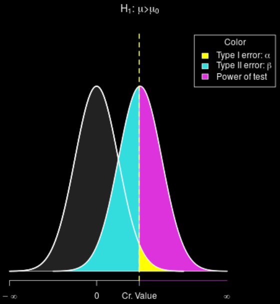
Two-sided test

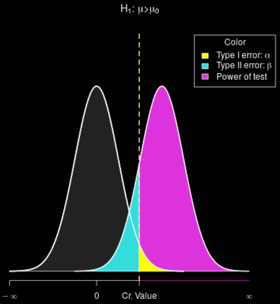


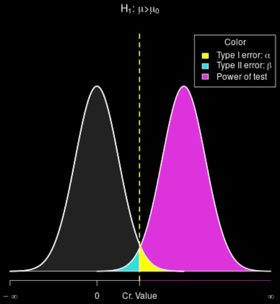




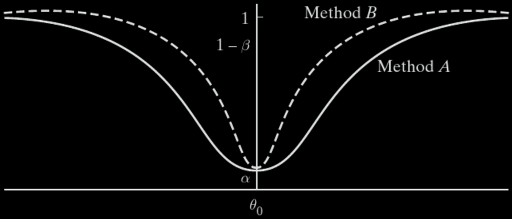
One-sided test



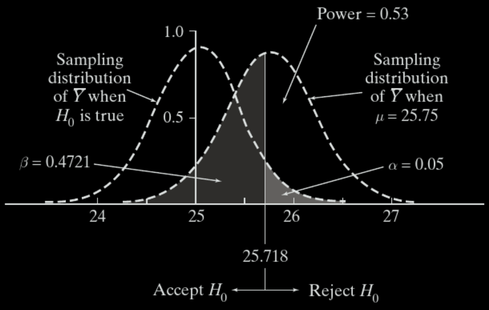


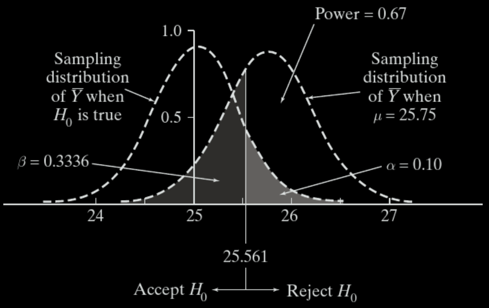


Use the **power curves** to select methods
(steepest one!)

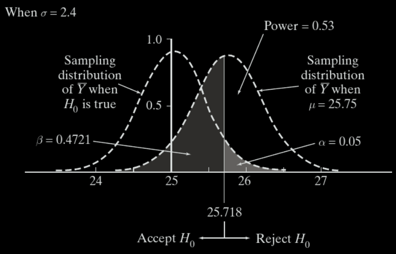


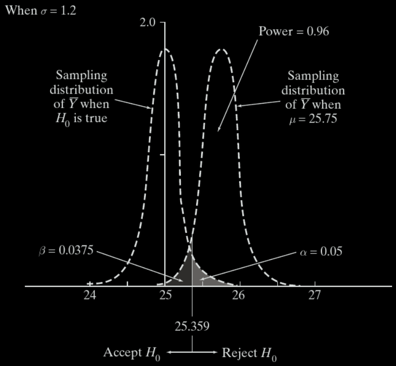
$$\alpha \uparrow \implies \beta \downarrow \quad \text{and} \quad (1-\beta) \uparrow$$





$$\sigma \downarrow \implies \beta \downarrow \quad \text{and} \quad (1-\beta) \uparrow$$





One usually cannot control the given parameter σ. But one can achieve the same power of test by increasing the sample size n.

E.g. Test $H_0 : \mu = 100$ v.s. $H_1 : \mu > 100$ at $\alpha = 0.05$ with $\sigma = 14$ known.
Requirement: $1 - \beta = 0.60$ when $\mu = 103$.
Find smallest sample size n.

Remark: Two condisions: $\alpha = 0.05$ and $1 - \beta = 0.60$
Two unknowns: Critical value y^* and sample size n

Sol.

$$C = \left\{ z : z = \frac{\bar{y} - \mu_0}{\sigma/\sqrt{n}} \geq z_\alpha \right\}.$$

$$
\begin{aligned}
1-\beta &= \mathbb{P}\left(\frac{\overline{Y}-\mu_0}{\sigma/\sqrt{n}} \geq z_\alpha \,\middle|\, \mu_1\right) \\
&= \mathbb{P}\left(\frac{\overline{Y}-\mu_1}{\sigma/\sqrt{n}} + \frac{\mu_1-\mu_0}{\sigma/\sqrt{n}} \geq z_\alpha \,\middle|\, \mu_1\right) \\
&= \mathbb{P}\left(Z \geq -\frac{\mu_1-\mu_0}{\sigma/\sqrt{n}} + z_\alpha \,\middle|\, \mu_1\right) \\
&= \Phi\left(\frac{\mu_1-\mu_0}{\sigma/\sqrt{n}} - z_\alpha\right)
\end{aligned}
$$

$$
\frac{\mu_1-\mu_0}{\sigma/\sqrt{n}} - z_\alpha = \Phi^{-1}(1-\beta) \quad \iff \quad n = \left(\sigma \times \frac{\Phi^{-1}(1-\beta)+z_\alpha}{\mu_1-\mu_0}\right)^2
$$

$$
n = \left\lceil \left(14 \times \frac{0.2533+1.645}{103-100}\right)^2 \right\rceil = \lceil 78.48 \rceil = 79.
$$

□

R	Python
$z_\alpha = \text{qnorm}(1-\alpha)$	$z_\alpha = \text{scipy.stats.norm.ppf}(1-\alpha)$
$\Phi^{-1}(1-\beta) = \text{qnorm}(1-\beta)$	$\Phi^{-1}(1-\beta) = \text{scipy.stats.norm.ppf}(1-\beta)$

Nonnormal data

Test $H_0 : \theta = \theta_0$, with $f_Y(y;\theta)$ is not normal distribution.

1. Identify a sufficient estimator $\widehat{\theta}$ for θ

2. Find the critical region C: Least compatible with H_0 but still admissible under H_1

3. Three types of questions:

   Given $\alpha \to$ find $C \to \beta$, $1-\beta$...

   From $C \to$ determine α

   From $\theta_e \to$ find P-value

Examples for nonnormal data

E.g. 1. A random sample of size n from uniform distr. $f_Y(y;\theta) = 1/\theta$, $y \in [0,\theta]$. To test

$$H_0 : \theta = 2.0 \quad \text{v.s.} \quad H_1 : \theta < 2.0$$

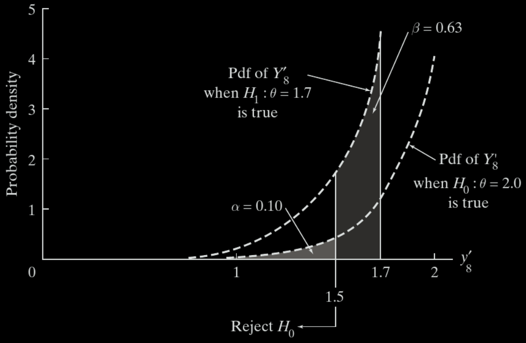
at the level $\alpha = 0.10$ of significance, one can use the decision rule based on Y_{max}. Find the probability of committing a Type II error when $\theta = 1.7$.

Remark: Y_{max} is a sufficient estimator for θ. Why?

Sol. 1) The critical region should has the form: $C = \{y_{max} : y_{max} \le c\}$.

2) We need to use the condition $\alpha = 0.10$ to find c.

3) Find the prob. of Type II error.



$$f_{Y_{max}}(y) = ... = n\frac{y^{n-1}}{\theta^n} \quad y \in [0, \theta].$$

$$\alpha = \int_0^c n\frac{y^{n-1}}{\theta_0^n}\mathrm{d}y = \left(\frac{c}{\theta_0}\right)^n \quad \Longrightarrow \quad c = \theta_0\alpha^{1/n} \qquad \text{(Under } H_0 : \theta = \theta_0\text{)}$$

$$\beta = \int_{\theta_0\alpha^{1/n}}^{\theta_1} n\frac{y^{n-1}}{\theta_1^n}\mathrm{d}y = 1 - \left(\frac{\theta_0}{\theta_1}\right)^n \alpha \qquad \text{(Under } \theta = \theta_1\text{)}$$

Finally, we need only plug in the values $\theta_0 = 2$, $\theta_1 = 1.7$ and $\alpha = 0.10$. □

E.g. 2. A random sample of size 4 from Poisson(λ): $p_X(k;\lambda) = e^{-\lambda}\lambda^k/k!$, $k = 0, 1, \cdots$. One wants to test

$$H_0 : \lambda = 0.8 \quad \text{v.s.} \quad H_1 : \lambda > 0.8.$$

at the level $\alpha = 0.10$. Find power of test when $\lambda = 1.2$.

Sol. 1) We've seen: $\overline{X} = \sum_{i=1}^{4} X_i$ is a sufficent estimator for λ;

$$\overline{X} \sim \text{Poisson}(3.2)$$

2) $C = \{\bar{k}; \bar{k} \geq c\}$.

3) $\alpha = 0.10 \to c = 6$.

4) Alternative $\lambda = 1.2 \to 1 - \beta = 0.35$.

Finding critical region

k	P(X=k)	P(X<= k)	P(X>k)	P(X>=k)
0	0.0408	0.0408	0.9592	1
1	0.1304	0.1712	0.8288	0.9592
2	0.2087	0.3799	0.6201	0.8288
3	0.2226	0.6025	0.3975	0.6201
4	0.1781	0.7806	0.2194	0.3975
5	0.114	0.8946	0.1054	0.2194
6	0.0608	0.9554	0.0446	0.1054
7	0.0278	0.9832	0.0168	0.0446
8	0.0111	0.9943	0.0057	0.0168
9	0.004	0.9982	0.0018	0.0057
10	0.0013	0.9995	0.0005	0.0018
11	0.0004	0.9999	0.0001	0.0005
12	0.0001	1	0	0.0001
13	0	1	0	0
14	0	1	0	0

Poisson lambda= 3.2

```
> qpois(1-0.10,3.2)
[1] 6
```

```
> scipy.stats.poisson.ppf(1-0.10,3.2)
[1] 6
```

Computing power of test

k	P(X=k)	P(X<= k)	P(X>k)	P(X>=k)
0	0.0082	0.0082	0.9918	1
1	0.0395	0.0477	0.9523	0.9918
2	0.0948	0.1425	0.8575	0.9523
3	0.1517	0.2942	0.7058	0.8575
4	0.182	0.4763	0.5237	0.7058
5	0.1747	0.651	0.349	0.5237
6	0.1398	0.7908	0.2092	0.349
7	0.0959	0.8867	0.1133	0.2092
8	0.0575	0.9442	0.0558	0.1133
9	0.0307	0.9749	0.0251	0.0558
10	0.0147	0.9896	0.0104	0.0251
11	0.0064	0.996	0.004	0.0104
12	0.0026	0.9986	0.0014	0.004
13	0.0009	0.9995	0.0005	0.0014
14	0.0003	0.9999	0.0001	0.0005
15	0.0001	1	0	0.0001
16	0	1	0	0
17	0	1	0	0
18	0	1	0	0
19	0	1	0	0
20	0	1	0	0

Poisson lambda= 4.8

$$1 - \beta = \mathbb{P}\left(\text{Reject } H_0 \mid H_1 \text{ is true}\right) = \mathbb{P}(\overline{X} \geq 6 | \overline{X} \sim \mathit{Poisson}(4.8))$$

□

```
> 1-ppois(6-1,4.8)
[1] 0.3489936
```

```
> 1-scipy.stats.poisson.cdf(6-1,4.8)
[1] 0.3489935627305083
```

```r
PlotPoissonTable <- function(n=14,lambda=3.2,png_filename,TableTitle) {
  library(gridExtra)
  library(grid)
  library(gtable)
  x = seq(1,n,1)
  # qpois(0.90,lambda)
  tb = cbind(x,
             round(dpois(x,lambda),4),
             round(ppois(x,lambda),4),
             round(1-ppois(x,lambda),4),
             round(c(1,(1-ppois(x,lambda))[1:n]),4))
  colnames(tb) <- c("k", "P(X=k)", "P(X<= k)", "P(X>k)", "P(X>=k)")
  rownames(tb) <-x
  table <- tableGrob(tb,rows = NULL)
  title <- textGrob(TableTitle,gp=gpar(fontsize=12))
  footnote <- textGrob(paste("Poisson lambda=",lambda),
                   x=0, hjust=0, gp=gpar( fontface="italic"))
  padding <- unit(0.2,"line")
  table <- gtable_add_rows(table, heights = grobHeight(title) + padding,pos = 0)
  table <- gtable_add_rows(table, heights = grobHeight(footnote)+ padding)
  table <- gtable_add_grob(table, list(title, footnote),
                         t=c(1, nrow(table)), l=c(1,2),r=ncol(table))
  png(png_filename)
  grid.draw(table)
  dev.off()
}

PlotPoissonTable(14,3.2,"Example_6-4-3_1.png","Finding critical region")
PlotPoissonTable(20,4.8,"Example_6-4-3_2.png","Computing power of test")
```

The *R* code to produce the previous two Poisson tables.

E.g. 3. A random sample of size 7 from $f_Y(y;\theta) = (\theta+1)y^\theta$, $y \in [0,1]$. Test

$$H_0 : \theta = 2.0 \quad \text{v.s.} \quad H_1 : \theta > 2.0$$

Decision rule: Let X be the number of y_i's that exceed 0.9;
Reject H_0 if $X \geq 4$.
Find α.

Sol. 1) $X \sim \text{binomial}(7, p)$.
2) Find p:

$$p = \mathbb{P}(Y \geq 0.9 | H_0 \text{ is true}) = \int_{0.9}^{1} 3y^2 \mathrm{d}y = 0.271$$

3) Compute α:

$$\alpha = \mathbb{P}(X \geq 4 | \theta = 2) = \sum_{k=4}^{7} \binom{7}{k} 0.271^k 0.729^{7-k} = 0.092.$$

□

```
> 1-pbinom(3,7,0.271)
[1] 0.09157663
```

```
> 1-scipy.stats.binom.cdf(3, 7, 0.271)
[1] 0.09157663095582469
```