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

§ 6.1 Introduction

§ 6.2 The Decision Rule

§ 6.3 Testing Binomial Data – $H_0 : p = p_0$

§ 6.4 Type I and Type II Errors

§ 6.5 A Notion of Optimality: The Generalized Likelihood Ratio

Instead of numerical estimates of parameters, in the form of either single points or confidence intervals, we want to make a choice between two conflicting theories, or **hypothesis**:

1. H_0: the null hypothesis

   v.s.

2. H_1: the alternative hypothesis

Comments: Hypothesis testing and confidence intervals are dual concepts to each other:

- One can be obtained from the other.
- However, it is often difficult to specify μ_0 to the null hypothesis.

Chapter 6. Hypothesis Testing

§ 6.1 Introduction

§ 6.2 The Decision Rule

§ 6.3 Testing Binomial Data – $H_0 : p = p_0$

§ 6.4 Type I and Type II Errors

§ 6.5 A Notion of Optimality: The Generalized Likelihood Ratio

Go over the example first....

Suppose our friend Jory claims that he has some magic power to predict the side of a randomly tossed fair-coin.

Jory claims that he could do more than
$1/2$
of the time on average.

Let's test Jory to see if we believe his claim.

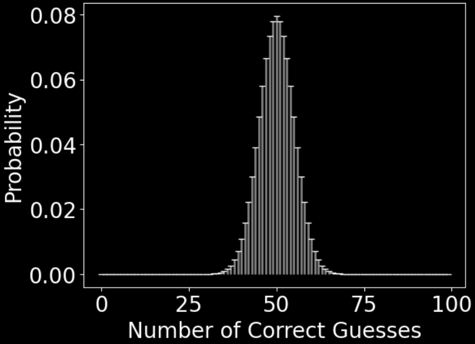
We made Jory guess a repeatedly tossed coin for 100 times.

He guesses correctly 54 times.

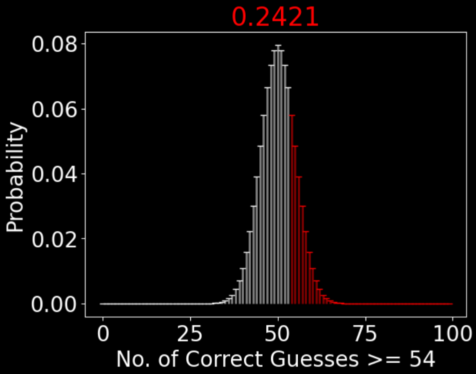
Question:

Does this provide strong evidence that Jory has the proclaimed magic power?

If Jory is guessing randomly, the number of correct guesses would follow a binomial distribution with parameters $n = 100$ and $p = 1/2$.



What is probability that Jory gets 54 or more correct when guessing randomly?



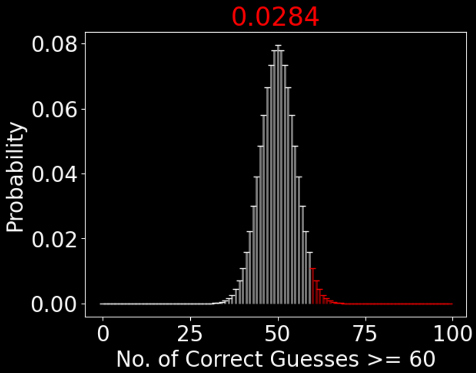
$$\mathbb{P}\left(X \geq 54\right) = \sum_{n=54}^{100} \binom{100}{n} \left(\frac{1}{2}\right)^{n} \left(\frac{1}{2}\right)^{100-n} = 0.2421.$$

It is not unlikely to get this many correct guesses due to chance.

Conclusion:

There is No strong evidence that Jory has better than a 1/2 chance of correctly guessing the coin.

What is probability that Jory gets 60 or more correct when guessing randomly?



$$\mathbb{P}\left(X \geq 60\right) = \sum_{n=60}^{100} \binom{100}{n} \left(\frac{1}{2}\right)^{n} \left(\frac{1}{2}\right)^{100-n} = 0.0284.$$

Either

Jory is purely guessing with probability of success of $\frac{1}{2}$, and we witnessed a very unusual event due to chance.

Or

Jory is truly having the magic power to guess the coin.

Conclusion:

We have strong evidence against
Red Hypothesis

Or the test is in favor of
Green Hypothesis

Before testing Jory, could you set up a threshold above which we will believe Jory's super power?

Find smallest m such that

$$\mathbb{P}\left(X \geq m\right) = \sum_{n=m}^{100} \binom{100}{n} \left(\frac{1}{2}\right)^{n} \left(\frac{1}{2}\right)^{100-n} \leq 0.05$$

$$\Downarrow$$

$$\boxed{m = 59}$$

b.c. $\mathbb{P}\left(X \geq 58\right) = 0.067$ & $\mathbb{P}\left(X \geq 59\right) = 0.044$

We have just informally conducted a hypothesis test with the null hypothesis

$$H_0 : p = \frac{1}{2}$$

against the alternative hypothesis

$$H_1 : p > \frac{1}{2}$$

under the significance level $\alpha = 0.05$

which is equivalent to either

producing the critical region $m \geq 59$ or comparing with the p-value.

- Test statistic: Any function of the observed data whose numerical value dictates whether H_0 is accepted or rejected.

- Critical region C: The set of values for the test statistic that result in the null hypothesis being rejected.

  Critical value: The particular point in C that separates the rejection region from the acceptance region.

- Level of significance α: The probability that the test statistic lies in the critical region C under H_0.

Test Normal mean $H_0 : \mu = \mu_0$ (σ known)

Setup:

1. Let $Y_1 = y_1, \cdots, Y_n = y_n$ be a random sample of size n from $N(\mu, \sigma^2)$ with σ known.
2. Set $\bar{y} = \frac{1}{n}(y_1 + \cdots + y_n)$ and $z = \frac{\bar{y} - \mu_0}{\sigma / \sqrt{n}}$.
3. The level of significance is α.

Test:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$$

reject H_0 if $z \geq z_\alpha$.

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$$

reject H_0 if $z \leq -z_\alpha$.

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

reject H_0 if $|z| \geq z_{\alpha/2}$.

- Simple hypothesis: Any hypothesis which specifies the population distribution completely.
- Composite hypothesis: Any hypothesis which does not specify the population distribution completely.

Conv. We always assume H_0 is simple and H_1 is composite.

Definition. The P-value associated with an observed test statistic is the probability of getting a value for that test statistic as extreme as or more extreme than what was actually observed (relative to H_1) given that H_0 is true.

Note: Test statistics that yield small P-values should be interpreted as evidence against H_0.

E.g. Suppose that test statistic $z = 0.60$. Find P-value for

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases} \qquad \begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases} \qquad \begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

$$\mathbb{P}(Z \geq 0.60) = 0.2743. \qquad \mathbb{P}(Z \leq 0.60) = 0.7257.$$

$$\begin{aligned} &\mathbb{P}(|Z| \geq 0.60) \\ &= 2 \times 0.2743 \\ &= 0.5486. \end{aligned}$$

Chapter 6. Hypothesis Testing

§ 6.1 Introduction

§ 6.2 The Decision Rule

§ 6.3 Testing Binomial Data – $H_0 : p = p_0$

§ 6.4 Type I and Type II Errors

§ 6.5 A Notion of Optimality: The Generalized Likelihood Ratio

Setup: Let $X_1 = k_1, \cdots, X_n = k_n$ be a‘ random sample of size n from Bernoulli(p). $X = \sum_{i=1}^{n} X_i \sim$ Binomial(n, p). We want to test $H_0 : p = p_0$.

1. When n is large, use Z score. Large-sample test
2. Otherwise, use the exact binomial distribution. Small-sample test

$$n \text{ is large}$$
$$\Updownarrow$$
$$0 < np_0 - 3\sqrt{np_0(1-p_0)} < np_0 + 3\sqrt{np_0(1-p_0)} < n$$
$$\Updownarrow$$
$$n > 9 \times \max\left(\frac{1-p_0}{p_0}, \frac{p_0}{1-p_0}\right).$$

Large-sample test for p

Setup:

1. Let $X_1 = k_1, \cdots, X_n = k_n$ be a random sample of size n from Bernoulli(p).
2. Suppose $n > 9 \max\left(\frac{1-p_0}{p_0}, \frac{p_0}{1-p_0}\right)$.
3. Set $k = k_1 + \cdots + k_n$ and $z = \frac{k - np_0}{\sqrt{np_0(1-p_0)}}$.
4. The level of significance is α.

Test:

$$\begin{cases} H_0 : p = p_0 \\ H_1 : p > p_0 \end{cases}$$

reject H_0 if $z \geq z_\alpha$.

$$\begin{cases} H_0 : p = p_0 \\ H_1 : p < p_0 \end{cases}$$

reject H_0 if $z \leq -z_\alpha$.

$$\begin{cases} H_0 : p = p_0 \\ H_1 : p \neq p_0 \end{cases}$$

reject H_0 if $|z| \geq z_{\alpha/2}$.

Small-sample test for p

E.g. $n = 19$, $p_0 = 0.85$, $\alpha = 0.10$. Find critical region for the two-sided test

$$\begin{cases} H_0 : p = p_0 \\ H_1 : p \neq p_0 \end{cases}$$

Sol. $19 = n < 9 \times \max\left(\frac{0.85}{0.15}, \frac{0.15}{0.85}\right) = 51$, so small sample test.

By checking the table, the critical region is

$$C = \{k : k \leq 13 \quad \text{or} \quad k = 19\},$$

so that

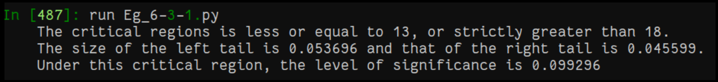
$$\begin{aligned} \alpha &= \mathbb{P}(X \in C | H_0 \text{ is true}) \\ &= \mathbb{P}(X \leq 13 | p = 0.85) + \mathbb{P}(X = 19 | p = 0.85) \\ &= 0.099295 \approx 0.10. \end{aligned}$$

$\square$

Binomial with n = 19 and p = 0.85

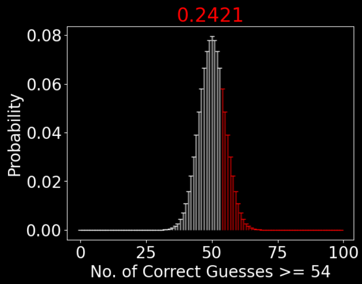
$$\begin{array}{rll}
x & P(X=x) & \\
\begin{array}{r}6\\7\\8\\9\\10\\11\\12\\13\end{array} & \left.\begin{array}{l}0.000000\\0.000002\\0.000018\\0.000123\\0.000699\\0.003242\\0.012246\\0.037366\end{array}\right\} & \rightarrow P(X \le 13) = 0.053696 \\
14 & 0.090746 & \\
15 & 0.171409 & \\
16 & 0.242829 & \\
17 & 0.242829 & \\
18 & 0.152892 & \\
19 & 0.045599 & \rightarrow P(X = 19) = 0.045599
\end{array}$$

```python
# Eg_6-3-1.py
from scipy.stats import binom
n = 19
p = 0.85
rv = binom(n, p)
low = rv.ppf(0.05)
upper = rv.ppf(0.95)
left  = round(rv.cdf(low), 6)
right = round(1-rv.cdf(upper), 6)
both = round(rv.cdf(low)+1-rv.cdf(upper), 6)
Results = """\
    The critical regions is less or equal to {low:.0f}, or strictly greater than {upper:.0f}.
    The size of the tail is {left:.6f} and that of the right tail is {right:.6f}.
    Under this critical region, the level of significance is {both:.6f}
""".format(**locals())
print(Results)
```



```
In [487]: run Eg_6-3-1.py
    The critical regions is less or equal to 13, or strictly greater than 18.
    The size of the left tail is 0.053696 and that of the right tail is 0.045599.
    Under this critical region, the level of significance is 0.099296
```

$$X \sim \text{Binomial}(100, 1/2)$$

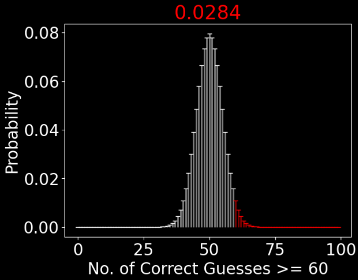


$$\mathbb{P}\left(X \geq 54\right) = \sum_{n=54}^{100} \binom{100}{n} \left(\frac{1}{2}\right)^{n} \left(\frac{1}{2}\right)^{100-n} = 0.2421.$$

vs

$$\mathbb{P}\left(\frac{X - 50}{\sqrt{100 \times \frac{1}{2} \times \frac{1}{2}}} \geq \frac{54 - 50}{\sqrt{100 \times \frac{1}{2} \times \frac{1}{2}}}\right) \approx \mathbb{P}\left(Z \geq \frac{4}{5}\right) = 0.2119$$

$$X \sim \text{Binomial}(100, 1/2)$$



$$\mathbb{P}\left(X \geq 60\right) = \sum_{n=60}^{100} \binom{100}{n} \left(\frac{1}{2}\right)^{n} \left(\frac{1}{2}\right)^{100-n} = 0.0284.$$

vs

$$\mathbb{P}\left(\frac{X - 50}{\sqrt{100 \times \frac{1}{2} \times \frac{1}{2}}} \geq \frac{60 - 50}{\sqrt{100 \times \frac{1}{2} \times \frac{1}{2}}}\right) \approx \mathbb{P}\left(Z \geq 2\right) = 0.0228$$

Chapter 6. Hypothesis Testing

§ 6.1 Introduction

§ 6.2 The Decision Rule

§ 6.3 Testing Binomial Data – $H_0 : p = p_0$

§ 6.4 Type I and Type II Errors

§ 6.5 A Notion of Optimality: The Generalized Likelihood Ratio

	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 of error types		**Null hypothesis (H_0) is**	
		True	**False**
Decision about null hypothesis (H_0)	**Don't reject**	Correct inference (true negative) (probability = 1 - α)	Type II error (false negative) (probability = β)
	Reject	Type I error (false positive) (probability = α)	Correct inference (true positive) (probability = 1 - β)

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.

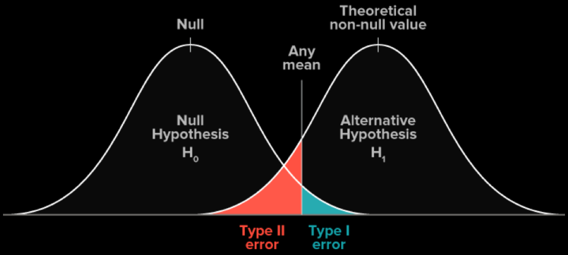


Figure: One-sided inference $H_1 : \mu > \mu_0$

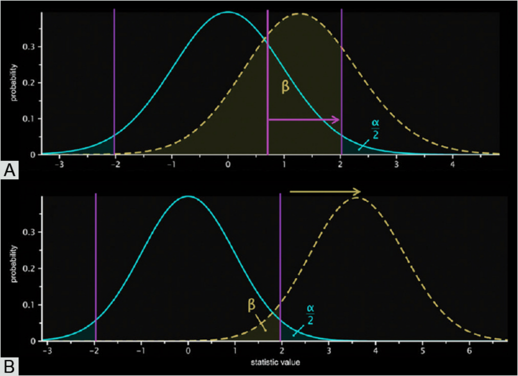
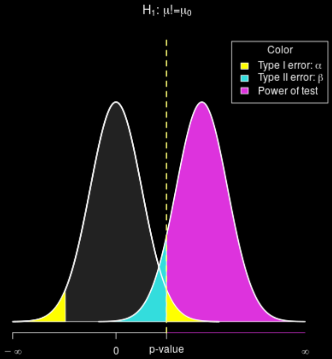
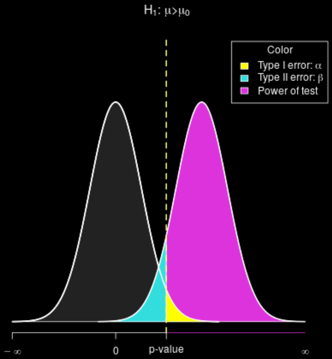


Figure: Two-sided inference $H_1 : \mu \neq \mu_0$

Power of test $1-\beta$

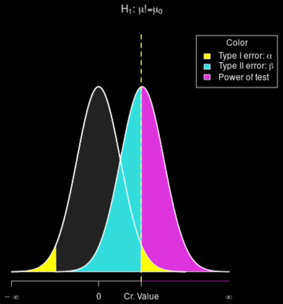
$$\text{Power of test} = \mathbb{P}(\text{Reject } H_0 | H_1 \text{ is true}) = 1-\beta$$

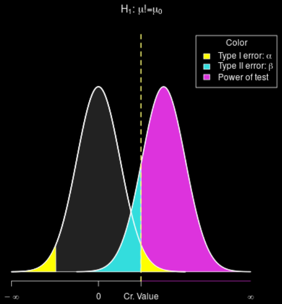


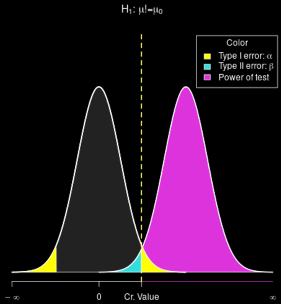


One online interactive show all α, β and $1-\beta$:
`https://rpsychologist.com/d3/NHST/`

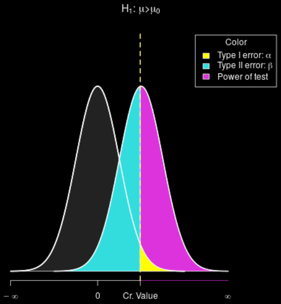
Two-sided test

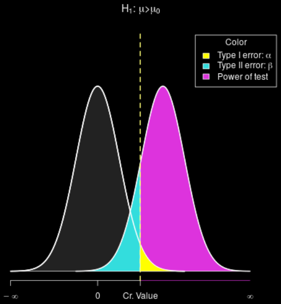


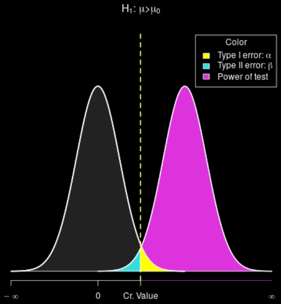




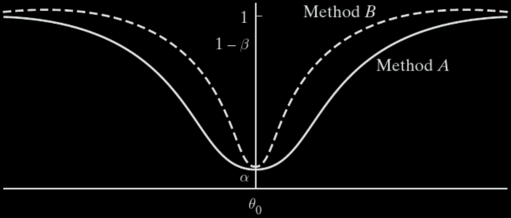
One-sided test



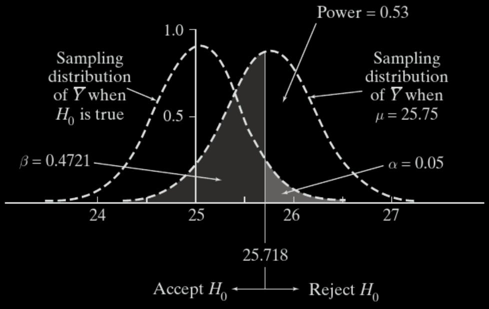


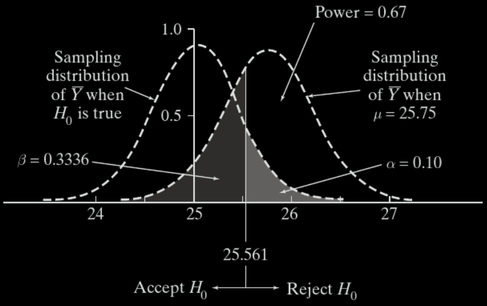


Use the **power curves** to select methods (steepest one!)



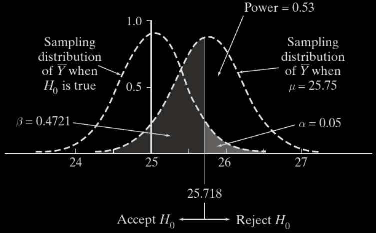
$$\alpha \uparrow \implies \beta \downarrow \quad \text{and} \quad (1-\beta) \uparrow$$



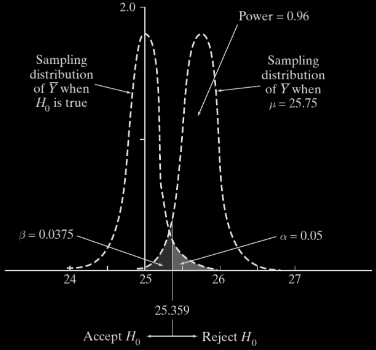


$$\sigma \downarrow \quad \Longrightarrow \quad \beta \downarrow \quad \text{and} \quad (1-\beta) \uparrow$$

When $\sigma = 2.4$



When $\sigma = 1.2$



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.

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

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

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

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

$\square$

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

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

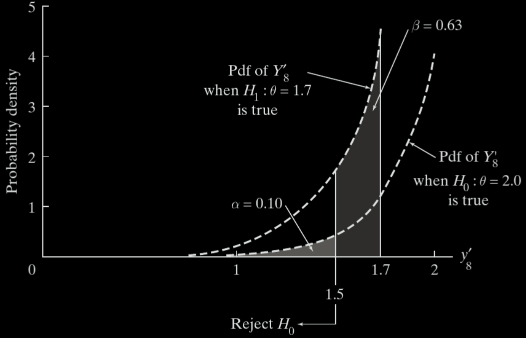
at the level $\alpha = 0.10$ of significance, one can use the decision rule based on Y_{max}. Find the probability of committing a Type II error when $\theta = 1.7$.

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

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

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

3) Find the prob. of Type II error.



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

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

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

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

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

$\square$

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

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

```
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:

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

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
```

Chapter 6. Hypothesis Testing

§ 6.1 Introduction

§ 6.2 The Decision Rule

§ 6.3 Testing Binomial Data – $H_0 : p = p_0$

§ 6.4 Type I and Type II Errors

§ 6.5 A Notion of Optimality: The Generalized Likelihood Ratio

Difficulties

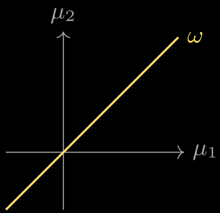
Scalar parameter		Vector parameter
Simple-vs-Composite test $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$	$\Rightarrow$	Composite-vs-Composite test $H_0 : \theta \in \omega$ vs $H_1 : \theta \in \Omega \cap \omega^c$

E.g. Two normal populations $N(\mu_i, \sigma_i)$, $i = 1, 2$. σ_i are known, μ_i unknown.

$$H_0 : \mu_1 = \mu_2 \quad \text{vs} \quad H_1 : \mu_1 \neq \mu_2.$$

Equivalently,

$$H_0 : (\mu_1, \mu_2) \in \omega \quad \text{vs} \quad H_1 : (\mu_1, \mu_2) \notin \omega.$$



- Let $Y_1, \cdots, Y_n$ be a random sample of size n from $f_Y(y; \theta_1, \cdots, \theta_k)$
- Let Ω be all possible values of the parameter vector $(\theta_1, \cdots, \theta_k)$
- Let $\omega \subseteq \Omega$ be a subset of Ω.

- Test:

$$H_0 : \theta \in \omega \quad \text{vs} \quad H_1 : \theta \in \Omega \setminus \omega.$$

- The **generalized likelihood ratio**, λ, is defined as

$$\lambda := \frac{\max_{(\theta_1, \cdots, \theta_k) \in \omega} L(\theta_1, \cdots, \theta_k)}{\max_{(\theta_1, \cdots, \theta_k) \in \Omega} L(\theta_1, \cdots, \theta_k)}$$

$$\lambda \in (0, 1]$$

λ close to zero
data NOT compatible with H_0
reject H_0

λ close to one
data compatible with H_0
accept H_0

- **Generalized likelihood ratio test (GLRT)**: Use the following critical region
$$C = \{\lambda : \lambda \in (0, \lambda^*]\}$$
to reject H_0 with either α or y^* being determined through
$$\alpha = \mathbb{P}\left(0 < \Lambda \leq \lambda^* \middle| H_0 \text{ is true}\right).$$

Remarks:

1. Maximization over Ω instead of $\Omega \setminus \omega$ in denominator:

   In practice, little effect on this change.

   In theory, much easier/nicer: $L(\theta_1, \cdots, \theta_k)$ is maximized over the whole space Ω by the max. likelihood estimates: $\Omega_e := (\theta_{e,1}, \cdots, \theta_{e,k}) \in \Omega$.

2. Suppose the maximization over ω is achieved at $\omega_e \in \omega$.

3. Hence:

$$\lambda = \frac{L(\omega_e)}{L(\Omega_e)}.$$

Remarks;

4. For simple-vs-composite test, $\omega = \{\omega_0\}$ consists only one point:

$$\lambda = \frac{L(\omega_0)}{L(\Omega_e)}.$$

5. Working with Λ is hard since $f_\Lambda(\lambda|H_0)$ is hard to obtain.

   If Λ is a *(monotonic) function* of some r.v. W, whose pdf is known.

   Suggesting testing procedure

   Test based on $\lambda \quad \Longleftrightarrow \quad$ Test based on w.

E.g. 1 Let $Y_1, \cdots, Y_n$ be a random sample of size n from the uniform pdf: $f_Y(y:\theta) = 1/\theta$, $y \in [0,\theta]$. Find the form of GLRT for

$$H_0 : \theta = \theta_0 \quad \text{v.s.} \quad H_1 : \theta < \theta_0 \qquad\qquad \text{with given } \alpha.$$

Sol. 1) The null hypothesis is simple, and hence

$$L(\omega_e) = L(\theta_0) = \theta_0^{-n} \prod_{i=1}^{n} I_{[0,\theta_0]}(y_i) = \theta^{-n} I_{[0,\theta_0]}(y_{max}).$$

2) The MLE for θ is y_{max} and hence,

$$L(\Omega_e) = L(y_{max}) = y_{max}^{-n} I_{[0,y_{max}]}(y_{max}) = y_{max}^{-n}.$$

3) Hence,

$$\lambda = \frac{L(\omega_e)}{L(\Omega_e)} = \left(\frac{y_{max}}{\theta_0}\right)^n I_{[0,\theta_0]}(y_{max})$$

that is, the test statistic is

$$\Lambda = \left(\frac{Y_{max}}{\theta_0}\right)^n I_{[0,\theta_0]}(Y_{max}).$$

4) α and critical value λ^*:

$$\begin{aligned}\alpha &= \mathbb{P}(0 < \Lambda \leq \lambda^* | H_0 \text{ is true}) \\ &= \mathbb{P}\left(\left[\frac{Y_{max}}{\theta_0}\right]^n I_{[0,\theta_0]}(Y_{max}) \leq \lambda^* \middle| H_0 \text{ is true}\right) \\ &= \mathbb{P}\left(Y_{max} \leq \theta_0 (\lambda^*)^{1/n} \middle| H_0 \text{ is true}\right)\end{aligned}$$

Λ suggests the test statistic Y_{max}:

$$\text{Test based on } \lambda \iff \text{Test based of } y_{max}$$

5) Let's find the pdf of Y_{max}. The cdf of Y is $F_Y(y;\theta_0) = y/\theta_0$ for $y \in [0,\theta_0]$. Hence,

$$\begin{aligned} f_{Y_{max}}(y;\theta_0) &= nF_Y(y;\theta_0)^{n-1} f_Y(y;\theta_0) \\ &= \frac{ny^{n-1}}{\theta_0^n}, \quad y \in [0,\theta_0]. \end{aligned}$$

6) Finally, by setting $y^* := \theta_0(\lambda^*)^{1/n}$, we see that

$$\begin{aligned} \alpha &= \mathbb{P}\left(Y_{max} \leq y^* \middle| H_0 \text{ is true}\right) \\ &= \int_0^{y^*} \frac{ny^{n-1}}{\theta_0^n} \mathrm{d}y \\ &= \frac{(y^*)^n}{\theta_0^n} \quad \iff \quad y^* = \theta_0 \alpha^{1/n}. \end{aligned}$$

7) Therefore, H_0 is rejected if

$$y_{max} \leq \theta_0 \alpha^{1/n}.$$

□

E.g. 2 Let $X_1, \cdots, X_n$ be a random sample from the geometric distribution with parameter p.

Find a test statistic Λ for testing $H_0 : p = p_0$ versus $H_1 : p \neq p_0$.

Sol. Let $\overline{X}$ and $\overline{k}$ be the sample mean. Because the null hypothesis is simple,

$$L(\omega_e) = L(p_0) = \prod_{i=1}^{n} (1 - p_0)^{k_i - 1} p_0 = (1 - p_0)^{n\bar{k} - n} p_0^n,$$

which shows that $\bar{k}$ is a sufficient estimator.

On the other hand, the MLE for the parameter p is $1/\bar{k}$. So

$$L(\Omega_e) = L(1/\bar{k}) = \prod_{i=1}^{n} \left(1 - \frac{1}{\bar{k}}\right)^{k_i - 1} \frac{1}{\bar{k}} = \left(\frac{\bar{k} - 1}{\bar{k}}\right)^{n\bar{k} - n} \frac{1}{\bar{k}^n}.$$

Hence,

$$\lambda = \frac{L(\omega_e)}{L(\Omega_e)} = \left(\frac{\bar{k}(1 - p_0)}{\bar{k} - 1}\right)^{n\bar{k} - n} (p_0 \bar{k})^n$$

Finally, $\Lambda = \left(\frac{\overline{X}(1-p_0)}{\overline{X}-1}\right)^{n\overline{X}-n} (p_0\overline{X})^n$. □

E.g. 3 Let $Y_1, \cdots, Y_n$ be a random sample from the exponential distribution with parameter λ.

Find a test statistic V for testing $H_0 : \lambda = \lambda_0$ versus $H_1 : \lambda \neq \lambda_0$.

Sol. Since the null hypothesis is simple,

$$L(\omega_e) = L(\lambda_0) = \prod_{i=1}^{n} \lambda_0 e^{-\lambda_0 y_i} = \lambda_0^n e^{-\lambda_0 \sum_{i=1}^n y_i}$$

Let $Z = \sum_{i=1}^n Y_i \sim \text{Gamma}(n, \lambda)$, which is a sufficient estimator.

On the other hand, the MLE for λ is $1/\bar{y} = n/z$:

$$L(\Omega_e) = L(1/\bar{y}) = (n/z)^n e^{-n}.$$

Hence,

$$\lambda = \frac{L(\omega_e)}{L(\Omega_e)} = z^n n^{-n} \lambda_0^n e^{-\lambda_0 z + n}$$

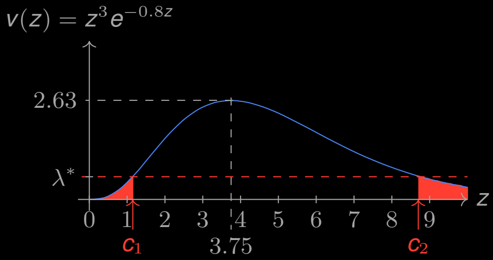
Finally, $\Lambda = Z^n n^{-n} \lambda_0^n e^{-\lambda_0 Z + n}$ or $V = Z^n e^{-\lambda_0 Z}$. □

The critical region in terms of V should be:

$$\begin{aligned} 0.05 = \alpha &= \mathbb{P}\left(V \in (0, y^*] \middle| H_0 \text{ is true} \right) \\ &= \int_0^{y^*} f_V(v) \mathrm{d}v \end{aligned}$$

However, it is not easy to find the exact distribution of V.

One can also make the inference based on the test statistic Z ...



This suggests that the critical region in terms of z should be of the form:

$$(0, c_1) \cup (c_2, \infty)$$

For convenience, we put $\alpha/2$ mass on each tails of the density of Z:

Find c_1 and c_2 such that

$$\int_0^{c_1} f_Z(z)dz = \int_{c_2}^{\infty} f_Z(z)dz = \frac{\alpha}{2}.$$

	using V	using Z
Critical region	$(0, v^*]$	$(0, z_1] \cup [z_2, \infty)$
pdf	hard to obtain	Gamma (n, λ)

E.g. 4 Let $Y_1, \cdots, Y_n$ be a random sample from $N(\mu, 1)$.
Find a test statistic Λ for testing $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$.

Sol. Since the null hypothesis is simple,

$$L(\omega_e) = L(\mu_0) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y_i-\mu_0)^2}{2}}.$$

On the other hand, the MLE for μ is $\bar{y}$:

$$L(\Omega_e) = L(\bar{y}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y_i-\bar{y})^2}{2}}.$$

Hence,

$$\lambda = \frac{L(\omega_e)}{L(\Omega_e)} = \exp\left(-\sum_{i=1}^{n} \frac{(y_i-\mu_0)^2 - (y_i-\bar{y})^2}{2}\right) = \exp\left(-\frac{n(\bar{y}-\mu_0)^2}{2}\right).$$

Finally, $\Lambda = \exp\left(-\frac{n}{2}\left(\overline{Y}-\mu_0\right)^2\right)$ or $V = \frac{\overline{Y}-\mu_0}{1/\sqrt{n}} \sim N(0,1)$ $\square$