Homework 6

Chapter 10.

p. 170-173.

Ex: 5, 7, 8a, 11a.

5. Let $X_1 \dots X_n \sim \text{Uniform}(0, \theta)$ and let $Y = \max\{X_1 \dots X_n\}$. We want to test

$$H_0: \theta = \frac{1}{2}$$
$$H_1: \theta > \frac{1}{2}$$

a) Find the power function.

$$\beta(\theta) = P_\theta(X \in R) = P(Y > c) =$$
$$= 1 - P(\max\{X_1 \dots X_n\} < c) =$$
$$= 1 - (P(X_1 < c) \cdot \dots \cdot P(X_n < c)) =$$
$$= 1 - \left(c \cdot \frac{1}{\theta}\right)^n = 1 - \left(\frac{c}{\theta}\right)^n$$

(HW 3, ex 7a)
(I did it here)

b) What choice of c will make the size of the test 0.05?

$$\beta\left(\frac{1}{2}\right) = 1 - \left(\frac{c}{\frac{1}{2}}\right)^n = 0.05$$

$$(2c)^n = 1 - 0.05$$

$$(2c)^n = 0.95$$

$$2c = \sqrt[n]{0.95}$$

$$c = \frac{0.95^{\left(\frac{1}{n}\right)}}{2}$$

c) In a sample of size $n = 20$ with $Y = 0.48$ what is the p-value? What conclusion about H_0 would you make?

$$p\text{-value} = P(Y > 0.48) = 1 - (2 \cdot 0.48)^{20} =$$
$$= 0.5579 \approx 0.5580$$

Conclusion: little or no evidence against H_0 $(0.5580 > 0.1)$

d) In a sample of size $n=20$ with $Y=0.52$ what is p-value? What conclusion about H_0 would you make?

$Y=0.52$ is already greater than $\frac{1}{2}$ by alternative hypothesis. ($H_0: \theta = \frac{1}{2}$).
It rejects the null hypothesis H_0. Therefore p-value is <0.01 very strong evidence against H_0. In this case p-value is 0.

7. From 8 Twain essays:

1) 0.225
2) 0.262
3) 0.217
4) 0.240
5) 0.230
6) 0.229
7) 0.235
8) 0.217

From 10 Snodgrass essay:

1) 0.209	6) 0.207
2) 0.205	7) 0.224
3) 0.196	8) 0.223
4) 0.210	9) 0.220
5) 0.202	10) 0.201

Comparing two means.

Let Twain = $T = \{X_1 \dots X_8\}$

Snodgrass = $S = \{Y_1 \dots Y_{10}\}$

Let's test null hypothesis $\mu_1 = \mu_2$

$H_0 : \delta = 0$

$H_1 : \delta \neq 0$, where $\delta = \mu_1 - \mu_2$

The non-parametric plug-in estimator is $\hat{\delta} = \overline{T} - \overline{S} = \overline{X} - \overline{Y}$

Estimated Standard Error

$$\hat{SE} = \sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}} = \sqrt{\frac{S_1^2}{8} + \frac{S_2^2}{10}}$$

The Wald test:

$$W = \frac{\hat{\theta} - \theta_0}{\hat{SE}} = \frac{\hat{\delta} - 0}{\hat{SE}} = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}}$$ (from book page 155)

$$\overline{X} = \frac{\sum_{i=1}^{8} X_i}{8} = 0.2319$$

$$\overline{Y} = \frac{\sum_{i=1}^{10} y_i}{10} = 0.2097$$

$$S_1^2 = \frac{\sum_{i=1}^{8} (0.2319 - X_i)^2}{7} = 0.00021$$

$$S_2^2 = \frac{\sum_{i=1}^{10} (0.2097 - y_i)^2}{9} = 0.9334 \cdot 10^{-4}$$

$$W = \frac{0.2319 - 0.2097}{\sqrt{\frac{0.00021}{8} + \frac{0.9334 \cdot 10^{-4}}{10}}} = 3.7036$$

Let w denote the observed value of the Wald statistic W. The p-value is given by

$$\text{p-value} = P_\theta(|W| > |w|) \approx P(|Z| > |w|) =$$
$$= 2\Phi(-|w|).$$ (page 158 of the book)

$$\text{p-value} = P(|Z| > |3.7036|) =$$
$$= P(Z > 3.7,\ Z < -3.7) = 2\Phi(-3.7) =$$
$$= 2 \cdot 0.0001078 \approx 0.0002$$

It is < 0.01, then very strong evidence against H_0

(Page 155 of book)

$$C = (\hat{\theta} - \hat{SE} \cdot z_{\alpha/2}, \ \hat{\theta} + \hat{SE} \cdot z_{\alpha/2})$$

$$C = (0.0222 - 0.0060 \cdot 1.96,$$
$$0.0222 + 0.0060 \cdot 1.96.)$$

$$C = (0.0104, \ 0.0339)$$

where

$$\hat{\theta} = \delta = \bar{X} - \bar{Y} = 0.022175$$

$$\hat{SE} = \sqrt{\frac{s_1^2}{8} + \frac{s_2^2}{10}} = 0.00599$$

$$z = 1.96$$

The size α Wald test reject H_0 when

$|W| > z_{\alpha/2}$

$|W| = 3.7$

$z_{\alpha/2} = 1.96 \Rightarrow 3.7 > 1.96$ (reject H_0)

I would say that theese are two different writers, because the means are different $\delta \neq 0$, where the H_0 is rejected. But the difference is small. (variance)

11Ex_Zhetessov.R × | 7_1_Zhetessov.R* ×

Source on Save | Run | Source

```r
#a) Wald Test
vec1 = c(.225, .262, .217, .240, .230, .229, .235, .217)
vec2 = c(.209, .205, .196, .210, .202, .207, .224, .223, .220, .201)

estimator =  mean(vec1)-mean(vec2)
SE = sqrt((var(vec1)/8)+(var(vec2)/10))

W = (estimator)/SE
W

CI= c(estimator - 1.96 * SE, estimator + 1.96 * SE)
CI

#b) Permutation Test
vec3 = c(vec1, vec2)
cnt <- 0

for(i in 1:1000000){
  x <- split(sample(vec3), rep(1:2, c(8,10)))
  res <- mean(x[[1]]) - mean(x[[2]])
  if(res > estimator){
    cnt <- cnt + 1
  }
}

p_value <- cnt / 1000000
p_value
#Zhetessov Nur
```

25:1 | (Top Level) | R Script

Console | Terminal | Jobs

R 4.1.2 · C:/Users/Nur/Desktop/дз/Стат/HW6/

```
> vec1 = c(.225, .262, .217, .240, .230, .229, .235, .217)
> vec2 = c(.209, .205, .196, .210, .202, .207, .224, .223, .220, .201)
> estimator =  mean(vec1)-mean(vec2)
> SE = sqrt((var(vec1)/8)+(var(vec2)/10))
> W = (estimator)/SE
> W
[1] 3.703554
> CI= c(estimator - 1.96 * SE, estimator + 1.96 * SE)
> CI
[1] 0.01043951 0.03391049
> vec3 = c(vec1, vec2)
> cnt <- 0
> for(i in 1:1000000){
+   x <- split(sample(vec3), rep(1:2, c(8,10)))
+   res <- mean(x[[1]]) - mean(x[[2]])
+   if(res > estimator){
+     cnt <- cnt + 1
+   }
+ }
> p_value <- cnt / 1000000
> p_value
[1] 0.00046
> #Zhetessov Nur
>
```

Environment | History | Connections | Tutorial

Import Dataset | 132 MiB | R | Global Environment

Data	
x	List of 2
Values	
CI	num [1:2] 0.0104 0.0339
cnt	460
estimator	0.022175
i	1000000L
p_value	0.00046
res	0.0134
SE	0.00598749275109745
vec1	num [1:8] 0.225 0.262 0.217 0.24 0.23 0.229 0.235 0.217
vec2	num [1:10] 0.209 0.205 0.196 0.21 0.202 0.207 0.224 0.223 0.22 0.201
vec3	num [1:18] 0.225 0.262 0.217 0.24 0.23 0.229 0.235 0.217 0.209 0.205 ...
W	3.70355354433382

Files | Plots | Packages | Help | Viewer

Zoom | Export

b) Now use a permutation test to avoid the use of large sample methods. What is your conclusion?

The p-value I got is between 0.004 and 0.005. The last test gave me 0.00483. Which 0.48%. The evidence is weak against H_0 of no difference.

8. Let $X_1, \dots, X_n \sim N(\theta, 1)$.

a) $H_0 : \theta = 0$

$H_1 : \theta = 1$ or $\theta \neq 0$

Power function

$$\beta(\theta) = P(T > c) = P(\underbrace{\sqrt{n}(T - \mu)}_{Z} > \sqrt{n}(c - \mu)) =$$

$$= P(Z > \sqrt{n}(c - \mu)) = 1 - \Phi(\sqrt{n}(c - \mu))$$

(From the lecture slides #13)

Significance level is α

$$\alpha = \max_{\theta \leq 0} \beta(\theta)$$

$\alpha = \beta(0) = 1 - \Phi(\sqrt{n}\cdot c)$

$\Phi(\sqrt{n}\cdot c) = 1 - \alpha$

$\sqrt{n}\cdot c = \Phi^{-1}(1-\alpha)$

$c = \frac{\Phi^{-1}(1-\alpha)}{\sqrt{n}}$

(Example 10.2, page 151)

11. A randomized, double-blind experiment has conducted to assess the effectiveness of several drugs for reducing postoperative nausea.

a) Test each drug vs placebo, at 5% level.

For each drug we do test

$H_0: p_1 = p_2$

$H_1: p_1 \neq p_2$

where p_1 - proportion of drug nausea

p_2 - proportion of placebo nausea

I will use the Wald test with proportions.

$$\widehat{SE} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{m} + \frac{\hat{p}_2(1-\hat{p}_2)}{n}}$$

$$W = \frac{\hat{\delta}-0}{\widehat{SE}} = \frac{\hat{p}_1-\hat{p}_2}{\widehat{SE}}$$

(Page 154. Example 10.7.)

Placebo $\hat{p}_2 = \frac{45}{80} = 0,5625$

	# total	# nausea	$\hat{p}_1$	$\hat{p}_1-\hat{p}_2$	$\widehat{SE}$	W	p-value
Chlorpromazine	75	26	0,3467	-0,2158	0.078	-2,76	0,0058
Dimenhydrinate	85	52	0,6118	0,0493	0.077	0.64	0,5222
Pentobarbital (100 mg)	67	35	0,5224	-0,0401	0.082	-0,49	0,6242
Pentobarbital (150 mg)	85	37	0,4353	-0,1272	0.077	-1.65	0,099

$$\text{p-value} = P(|Z|>|w|) = 2\Phi(-|w|)$$

(Theorem 10.13 page 158)

Also report the estimated odds-ratios.

The odd-ratio - is calculated by dividing the odds of the first group by the odds of the second group.

Odd - the ratio of number of successes by the number of failures.

Placebo odd $= \frac{45}{35}$ Odd-ratio $= \frac{\text{placebo-odd}}{\text{drug-odd}}$

	Success	Failure	Odd	Odd-ratio
Chlorpromazine	26	49	0,5306	2,42
Dimenhydrinate	52	33	1,576	0,816
Pentobarbital (100 mg)	35	32	1,093	1,176
Pentobarbital (150 mg)	37	48	0,77	1,67

Reject H_0 if $|W| > z_{\alpha/2}$

at 5% sign. level z is 1.96.

All the drug could not reject the null hypothesis about no effect. Except Chlorpromazine, whose $|W| = 2{,}76$, which is greater than $z = 1{,}96$. So, I conclude Chlorpromazine have significant results. p-value also gives strong evidence that this drug has difference effect than placebo, $(0{,}0058 < 0.01)$, or strong evidence against H_0.

7_1_Zhetessov.R | 11Ex_Zhetessov.R

```r
exp.data <- data.frame(
  name = c ("Placebo", "Chlorpromazine", "Dimenhydrinate", "Pentobarbital (100 mg)", "Pentobarbital (150 mg)"),
  total = c(80, 75, 85, 67, 85),
  nausea = c(45, 26, 52, 35, 37)
)

p_hat <- exp.data$nausea/exp.data$total #Вычисляю пропорции
p_hat

delta_hat <- p_hat[2:5] - p_hat[1] #Вычисляю estimator дельта_hat (Лекарства - Плацебо)
delta_hat

SE_hat <- sqrt( (p_hat[1]*(1-p_hat[1])/exp.data[1,]$total) + (p_hat[2:5]*(1-p_hat[2:5])/exp.data[c(2:5),]$total) )
#Вычисляю стандартное отклонение для всех лекарств
SE_hat

W <- delta_hat / SE_hat
W

p_values <- 2 * pnorm(-abs(W))
p_values

reject_H0 <- c()
for(i in W){
  if(abs(i)>1.96){
    reject_H0 <- append(reject_H0, i)
  }
}
```

Console | Terminal | Jobs

R 4.1.2 · ~/

```
+ )
> p_hat <- exp.data$nausea/exp.data$total #Вычисляю пропорции
> p_hat
[1] 0.5625000 0.3466667 0.6117647 0.5223881 0.4352941
> delta_hat <- p_hat[2:5] - p_hat[1] #Вычисляю estimator дельта_hat (Лекарства - Плацебо)
> delta_hat
[1] -0.21583333  0.04926471 -0.04011194 -0.12720588
> SE_hat <- sqrt( (p_hat[1]*(1-p_hat[1])/exp.data[1,]$total) + (p_hat[2:5]*(1-p_hat[2:5])/exp.data[c(2:5),]$total) )
> #Вычисляю стандартное отклонение для всех лекарств
> SE_hat
[1] 0.07807704 0.07661848 0.08246232 0.07725342
> W <- delta_hat / SE_hat
> W
[1] -2.7643638  0.6429873 -0.4864275 -1.6466051
> p_values <- 2 * pnorm(-abs(W))
> p_values
[1] 0.005703392 0.520232365 0.626664095 0.099639233
> reject_H0 <- c()
> for(i in W){
+   if(abs(i)>1.96){
+     reject_H0 <- append(reject_H0, i)
+   }
+ }
> reject_H0
[1] -2.764364
> #Zhetessov Nur
```