

Quiz 4 Notes Coursera Statistical Inference

Question 1 Hypothesis testing

In a random sample of 100 patients at a clinic, you would like to test whether the mean RDI is x or more using a one sided 5% type 1 error rate. The sample mean RDI has a mean of 12 and a standard deviation of 4. What is the largest value of x ($H_0: \mu = x$ versus $H_A: \mu > x$) would you reject for? This is a confidence interval question

- Choices:
 - A = 11.34
 - B = 10.84
 - C = 12.66
 - D = 9.30

We are looking for the smallest pvalue to give us the answer

```
n = 100
m = 12
z = qnorm(0.05)
sd = 4
alpha = .05
mu0 = m + z *sd /sqrt(n)
mu0
```

[1] 11.34

The result of 11.34 is the largest value we can use given the null and alternative hypothesis pair with a z value associated with 0.05.

Question 2 T test

A pharmaceutical company is interested in testing a potential blood pressure lowering medication. Their first examination considers only subjects that received the medication at baseline then two weeks later. The data are as follows (SBP in mmHg).

This questions references the one sample, two sided t test find the pvalue.

```
pharm <- data.frame(baseline=c(140,138,150,148,135),
week2=c(132,135,151,146,130))
t.test(pharm$baseline,pharm$week2, alternative =
"two.sided", paired=TRUE)
```

```
##
## Paired t-test
##
## data: pharm$baseline and pharm$week2
## t = 2.262, df = 4, p-value = 0.08652
## alternative hypothesis: true difference in means is not
## equal to 0
## 95 percent confidence interval:
## -0.7739 7.5739
## sample estimates:
## mean of the differences
## 3.4
```

The answer is p-value = 0.08652

Question 3 Confidence intervals

Test the hypothesis that $H_0: \mu = \mu_0$ and $H_A: \mu \neq \mu_0$. Two sided t-test. Alpha = .05.

```
n = 9
x_bar = 1100
sdev = 30
a = 0.05
round(x_bar + c(-1,1)*qt(1 - (a/2), n-1) * sdev/sqrt(n))
```

```
## [1] 1077 1123
```

The range of values for this is 1077 to 1123.

Question 4 P-value

3 out of 4 people chose Coke. Report a p-value for a test of the hypothesis that Coke is preferred to Pepsi using a one-sided exact test.

```
library(stats)
binom.test(x = 3, n = 4, p=.5, alternative="greater")
```

```
##
## Exact binomial test
##
## data: 3 and 4
## number of successes = 3, number of trials = 4, p-value =
0.3125
## alternative hypothesis: true probability of success is
greater than 0.5
## 95 percent confidence interval:
## 0.2486 1.0000
## sample estimates:
## probability of success
## 0.75
```

p-value = 0.3125

Question 5 Infection rate for a hospital

Find the one sided p-value for the relevant test of whether the hospital is *below* the standard.

```
p = 1/100
p_ = 10/1787
n=1787
serror = sqrt(p*(1-p)/n)
test_z = (p-p_)/serror

pnorm(test_z, lower.tail=FALSE)
```

```
## [1] 0.03067
```

The pvalue is: 0.03067

Question 6 Pvalue for two sided t-test

Number of total subjects = 18 diet = 9 and placebo = 9 Treated (mean = -3kg/m², st.dev = 1.5kg/m²)

Placebo (mean = 1kg/m², st.dev = 1.8kg/m²)

Find the pvalue for a two sided t test = less than 0.01

```
mean.diff = -3-1
df = (9 + 9 - 2)
m_tr = -3
m_pb = 1
s_tr = 1.5
s_pb = 1.8
pooled.var = (s_tr^2 * 9 + s_pb^2 * 9)/df
se.diff = sqrt(pooled.var/9 + pooled.var/9)
t.obt = mean.diff / se.diff
t.obt
```

```
## [1] -4.829
```

```
p.value = 2*pt(t.obt, df=df) #two tailed
p.value
```

```
## [1] 0.0001852
```

Question 7 Hypothesis test

sample: 9 men 90% confidence interval of 1,077 cc to 1,123cc. Would you reject a two sided 5% hypothesis test of $H_0: \mu = 1,078$ with the $H_A: \mu \neq 1,078$.

The main thing to focus on here is that we are give then 90% confidence intervals and not the 95% confidence interval. Using a two sided hypothesis test.

This is a two sided hypothesis test because of the \neq sign Under $H_0 \bar{x} = N(0.05, \sqrt{n})$ with $N(\text{volume loss}, .05/\sqrt{n})$

Now, we don't have to do any calculations because the 95% confidence region contains the 90% confidence interval. We also know that since $H_0: \mu = 1,078$ falls within the region it is plausible and we cannot reject the null because 1,078 is a plausible value of the population parameter.

Question 8 Power

Given Information:

- wanted sample size = 100 healthy adults

- why: to detect a four year mean brain volume loss of 0.01mm^3 .
- assume: st.dev for a four year loss in population = 0.04mm^3
- $\alpha = 0.05$ Find: the power of the study for a 5% one sided test versus a null hypothesis of no volume loss.

```
power.t.test(n=100, delta = .01, sd = .04, type =
"one.sample", alt = "one.sided")$power
```

```
## [1] 0.799
```

The power would be roughly 0.80

Question 9 find n for 90% power of type one error rate of 5%

Given:

- four year mean brain volume loss = $.01\text{mm}^3$
- $\text{sd} = .04\text{mm}^3$
- 90% power
- error rate of 5%
- one sided test versus a null hypothesis of no volume loss

Wanted: n

```
power.t.test(power = .9, delta = .01, sd = .04,
type="one.sample", alt="one.sided")$n
```

```
## [1] 138.4
```

Question 10 Power and alpha

As you increase the type one error rate, α , power increases, or larger power.

Question 11 Hypothesis testing and kryptonite

Given Group 1:

- sample = 288
- mean = 44 ll/ml
- sd = 12

Given Group 2:

- sample = 288
- mean = 42.04
- sd = 12

Wanted, p-value for a two sided z test of the relevant hypothesis.

```
mean.1 = 44
mean.2 = 42.04
sds = 12
n = 288

st.error = sds*sqrt(1/n + 1/n)
z.11 = (mean.1 - mean.2)/st.error
z.11
```

```
## [1] 1.96
```

```
p.v = 2*pnorm(-abs(z.11))
p.v
```

```
## [1] 0.05
```

The p.value is 0.05

Question 12

Suppose that a researcher performs 10 hypothesis tests and wants a familywise error rate of no more than 5%. What alpha level should she compare her pvalues to in order to ensure the desired error rate? Express as a proportion to two decimal places. The equation for the FER $P(\text{at least one significant result})$ is $1 - (1 - \alpha)^C$. Then we have $P(\text{at least one significant result}) = 1 - P(\text{no significant results})$

```
FER = 0.05/10
FER
```

```
## [1] 0.005
```