# Homework 4 for Stat Inference

*Brian Caffo*

## About these slides

- These are some practice problems for Statistical Inference Quiz 4
- They were created using slidify interactive which you will learn in Creating Data Products
- Please help improve this with pull requests here (https://github.com/bcaffo/courses)

Load the data set `mtcars` in the `datasets` R package. Assume that the data set mtcars is a random sample. Compute the mean MPG, $\bar{x}$, of this sample.

You want to test whether the true MPG is $\mu_0$ or smaller using a one sided 5% level test. ($H_0 : \mu = \mu_0$ versus $H_a : \mu < \mu_0$). Using that data set and a Z test:

1. . Based on the mean MPG of the sample $\bar{x}$, and by using a Z test: what is the smallest value of $\mu_0$ that you would reject for (to two decimal places)?

*** .hint This is the inversion of a one sided hypothesis test. It yields confidence bounds. (Note inverting a two sidded test yields confidence intervals.) Think about the derivation of the confidence interval.

*** .explanation We want to solve

$$\frac{\sqrt{n}(\bar{X} - \mu_0)}{s} = Z_{0.05}$$

Or

$$\mu_0 = \bar{X} - Z_{0.05} s/\sqrt{n} = \bar{X} + Z_{0.95} s/\sqrt{n}$$

Note that the quantile is negative.

```
mn <- mean(mtcars$mpg)
s <- sd(mtcars$mpg)
z <- qnorm(.05)
mu0 <- mn - z * s / sqrt(nrow(mtcars))
```

Note, it's easy to get tripped up in this problem on signs. If you get a value that's less than $\bar{X}$, then clearly it's wrong, since you'd never reject for $H_0 : \mu = \mu_0$ versus $H_a : \mu < \mu_0$ if $\mu_0$ was less than your observed mean. Also note the answer to "What is the largest value for which you would reject for?" is infinity.

21.84

Consider again the `mtcars` dataset. Use a two group t-test to test the hypothesis that the 4 and 6 cyl cars have the same mpg. Use a two sided test with unequal variances.

1. Do you reject at the 5% level (enter 0 for no, 1 for yes).
2. What is the P-value to 4 decimal places expressed as a proportion?

*** .hint Use `t.test` with the options `var.equal=FALSE`, `paired=FALSE`, `altnernative` as `two.sided`.

*** .explanation

```r
m4 <- mtcars$mpg[mtcars$cyl == 4]
m6 <- mtcars$mpg[mtcars$cyl == 6]
p <- t.test(m4, m6, paired = FALSE, alternative="two.sided", var.equal=FALSE)$p.value
```

The answer to 1 is 1 The answer to 2 is 4e-04

A sample of 100 men yielded an average PSA level of 3.0 with a sd of 1.1. What are the complete set of values that a 5% two sided Z test of $H_0 : \mu = \mu_0$ would fail to reject the null hypothesis for?

1. Enter the lower value to 2 decimal places.
2. Enter the upper value to 2 decimal places.

*** .hint This is equivalent to the confidence interval.

*** .explanation The answer to 1 is 2.78 The answer to 2 is 3.22

You believe the coin that you're flipping is biased towards heads. You get 55 heads out of 100 flips.

1. What's the exact relevant pvalue to 4 decimal places (expressed as a proportion)?
2. Would you reject a 1 sided hypothesis at $\alpha = .05$? (0 for no 1 for yes)?

*** .hint Use `pbinom` for a hypothesis that $p = .5$ veruss $p > .5$ where $p$ is the binomial success probability.

*** .explanation Note you have to start at 54 as it `lower.tail = FALSE` gives the strictly greater than probabilities

```r
ans <- round(pbinom(54, prob = .5, size = 100, lower.tail = FALSE), 4)
```

The answer to 1 is 0.1841 The answer to 2 is 0

A web site was monitored for a year and it received 520 hits per day. In the first 30 days in the next year, the site received 15,800 hits. Assuming that web hits are Poisson.

1. Give an exact one sided P-value to the hypothesis that web hits are up this year over last to four significant digits (expressed as a proportion).
2. Does the one sided test reject (0 for no 1 for yes)?

*** .hint Consider using `ppois` with $\lambda = 520 * 30$. Note this is nearly exactly Gaussian, so one could get away with the Gaussian calculation.

*** .explanation This test comes with the important assumption that web hits are a Poisson process.

```r
pv <- ppois(15800 - 1, lambda = 520 * 30, lower.tail = FALSE)
```

The answer to 1 is 0.0553 The answer to 2 is 0

Also, compare with the Gaussian approximation where $\hat{\lambda} \sim N(\lambda, \lambda/t)$

```r
pnorm(15800 / 30, mean = 520, sd = sqrt(520 / 30), lower.tail = FALSE)
```

```
## [1] 0.05465729
```

As $t \to \infty$ this becomes more Gaussian. The approximation is pretty good for this setting.

Suppose that in an AB test, one advertising scheme led to an average of 10 purchases per day for a sample of 100 days, while the other led to 11 purchaces per day, also for a sample of 100 days. Assuming a common standard deviation of 4 purchases per day. Assuming that the groups are independent and that they days are iid, perform a Z test of equivalence.

1. What is the P-value reported to 3 digits expressed as a proportion?
2. Do you reject the test? (0 for no 1 for yes).

*** .hint The standard error is

$$s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

*** .explanation

```
m1 <- 10; m2 <- 11
n1 <- n2 <- 100
s <- 4
se <- s * sqrt(1 / n1 + 1 / n2)
ts <- (m2 - m1) / se
pv <- 2 * pnorm(-abs(ts))
```

The answer to 1 is 0.077 The answer to 2 is 0

A confidence interval for the mean contains:

1. *All of the values of the hypothesized mean for which we would fail to reject with* $\alpha = 1 - Conf.Level$.
2. All of the values of the hypothesized mean for which we would fail to reject with $2\alpha = 1 - Conf.Level$.
3. All of the values of the hypothesized mean for which we would reject with $\alpha = 1 - Conf.Level$.
4. All of the values of the hypothesized mean for which we would reject with $2\alpha = 1 - Conf.Level$.

*** .hint This is directly from the notes. Note that a confidence interval gives values of $\mu$ that are supported by the data whereas a test rejects for values of $\mu$ different from $\mu_0$.

*** .explanation The only complicated part of this is the 2. Note that a 95% interval corresponds to a 5% level two sided test. So it's $\alpha = 1 - Conf.Level$. The confusion is that for both the two sided test and confidence interval, one needs to calculate $Z_{1-\alpha/2}$ (or the relevant T quantile).

Consider two problems previous. Assuming that 10 purchases per day is a benchmark null value, that days are iid and that the standard deviation is 4 purchases for day. Suppose that you plan on sampling 100 days. What would be the power for a one sided 5% Z mean test that purchases per day have increased under the alternative of $\mu = 11$ purchase per day?

1. Give your result as a proportion to 3 decimal places.

*** .hint Under $H_0$ $\bar X \sim N(10, .4)$. Under $H_a$ $\bar X \sim N(11, .4)$. We reject when $\bar X \geq 10 + Z_{.95} .4$.

*** .explanation The hint prettty much gives it away.

```
power <- pnorm(10 + qnorm(.95) * .4, mean = 11, sd = .4, lower.tail = FALSE)
```

The answer is 0.804

Researchers would like to conduct a study of healthy adults to detect a four year mean brain volume loss of .01 mm3. Assume that the standard deviation of four year volume loss in this population is .04 mm3.

1. What is necessary sample size for the study for a 5% one sided test versus a null hypothesis of no volume loss to achieve 80% power? (Always round up)

*** .hint Under $H_0$ $\bar{X}$ is $N(0, .05/\sqrt{n})$ and is $N(.01, .05/\sqrt{n})$ under $H_a$. We will reject if

$$\bar{X} \geq Z_{.95} s/sqrtn$$

which has probability 0.05 under $H_0$. Under $H_a$ it has probability

$$P\left(\frac{\bar{X} - 0.01}{s/\sqrt{n}} \geq \frac{.01}{s/\sqrt{n}} + z_{.95}\right) = P\left(Z \geq \frac{.01}{s/\sqrt{n}} + z_{.95}\right)$$

*** .explanation Looking at the hint we set

$$\frac{.01}{s/\sqrt{n}} + z_{.95} = z_{.2}$$

$$n = \frac{(z_{.95} - z_{.2})^2 s^2}{.01^2} = \frac{(z_{.95} + z_{.8})^2 s^2}{.01^2}$$

So we get

```
n <- (qnorm(.95) + qnorm(.8)) ^ 2 * .04 ^ 2 / .01^2
```

The answer is 99

In a court of law, all things being equal, if via policy you require a lower standard of evidence to convict people then

1. Less guilty people will be convicted.
2. *More innocent people will be convicted.*
3. More Innocent people will be not convicted.

*** .hint Think about it.

*** .explanation If you require less evidence to convict, then you will convict more people, guilty and innocent. Relate this property back to hypothesis tests.

Consider the `mtcars` data set.

1. Give the p-value for a t-test comparing MPG for 6 and 8 cylinder cars assuming equal variance, as a proportion to 3 decimal places.
2. Give the associated P-value for a z test.
3. Give the common (pooled) standard deviation estimate for MPG across cylinders to 3 decimal places.
4. Would the t test reject at the two sided 0.05 level (0 for no 1 for yes)?

*** .hint

```
mpg8 <- mtcars$mpg[mtcars$cyl == 8]
mpg6 <- mtcars$mpg[mtcars$cyl == 6]
m8 <- mean(mpg8); m6 <- mean(mpg6)
s8 <- sd(mpg8); s6 <- sd(mpg6)
n8 <- length(mpg8); n6 <- length(mpg6)
```

*** .explanation

```r
p <- t.test(mpg8, mpg6, paired = FALSE, alternative="two.sided", var.equal=TRUE)$p.value
mixprob <- (n8 - 1) / (n8 + n6 - 2)
s <- sqrt(mixprob * s8 ^ 2  +  (1 - mixprob) * s6 ^ 2)
z <- (m8 - m6) / (s * sqrt(1 / n8 + 1 / n6))
pz <- 2 * pnorm(-abs(z))
## Hand calculating the T just to check
#2 * pt(-abs(z), df = n8 + n6 - 2)
```

1. 0
2. 0
3. 2.27
4. 1

The Bonferonni correction controls this

1. False discovery rate
2. *The familywise error rate*
3. The rate of true rejections
4. The number of true rejections

*** .hint This is pretty much straight out of the notes

*** .explanation The Bonferonni correction is the classic correction for the familywise error rate.