

# Homework 8

## Significance Testing

Jaiden Atterbury

---

### Instructions

Please answer the following questions in the order in which they are posed. Add a few empty lines below each and write your answers there. **Focus on answering in complete sentences and show work whether we ask for it or not.** You will also need scratch paper/pen to work out the answers before typing it.

For help with formatting documents in RMarkdown, please consult R Markdown: The Definitive Guide. Another option is to search using Google.

---

### Exercises

1. (NHL) After a 2010 NHL playoff win in which Detroit Red-Wings wingman Henrik Zetterberg scored two goals in a 3-0 win over the Phoenix Coyotes. Detroit coach Mike Babcock said "He's been really good at playoff time each and every year. *He seems to score at a higher rate in playoffs compared to the regular season.*"

In 506 **regular season games**, Zetterberg scored 206 goals (goal scoring rate of  $\frac{206}{506} = 0.407$ ). In 89 **playoff games**, he scored 44 goals (goal scoring rate of  $\frac{44}{89} = 0.494$ ). Clearly, he has a higher goal scoring rate in the playoffs, but can it be explained by the vagaries of random chance? Or is the difference statistically significant?

Let  $X$  denote the number of goals he scores in the 89 playoff games. Assume

$$X \sim \text{Pois}(89 \times \lambda_0)$$

where  $\lambda_0$  is his goal scoring rate per game during the playoffs.

- a. State the null and alternative hypothesis.

Let  $X$  denote the number of goals scored by Henrik Zetterberg in 89 playoff games. It follows that  $X \sim \text{Pois}(89 \times \lambda_0)$  where  $\lambda_0$  is his goal scoring rate per game during the playoffs.

Since Henrik Zetterberg has a goal scoring rate of  $\frac{206}{506} = 0.407$  in regular season games, and the null hypothesis corresponds in some sense to no effect, our null hypothesis will be  $H_0 : \lambda_0 = \frac{206}{506}$ . Furthermore, since Detroit coach Mike Babcock said "*He seems to score at a higher rate in playoffs compared to the regular season.*", we know that our alternative hypothesis won't be two-sided. Instead, our alternative hypothesis will be  $H_1 : \lambda_0 > \frac{206}{506}$ .

To state it more formally, our null and alternative hypotheses will be:

$$H_0 : \lambda_0 = \frac{206}{506} \quad H_1 : \lambda_0 > \frac{206}{506}$$

- b. Calculate the P-value and summarize what you learn from it. Show your code for calculating the P-value. (You do not need to show the entire Poisson distribution)

In order to calculate the P-value and summarize what we learn from it, we will need to find the sampling distribution of our test statistic under the null hypothesis, as well as figure out which tail supports  $H_1$ . Under  $H_0$ , the sampling distribution of our test statistic is,  $X \sim \text{Pois}(89 \times \frac{206}{506})$ . Since  $H_1$  corresponds to an increasing rate of scoring, and  $E[X] = \lambda_0$ , we will expect the amount of goals scored to increase with increasing  $\lambda_0$ , thus we will look at the upper tail in this case.

Since Henrik has scored 44 goals in 89 playoff appearances, in order to find the P-value, we must find the probability associated with the following probability statement:  $P(X \geq 44)$ . We will calculate this in the following R chunk.

```
pval_1b <- ppois(q = 43, lambda = 89 * 206/506, lower.tail = F)
```

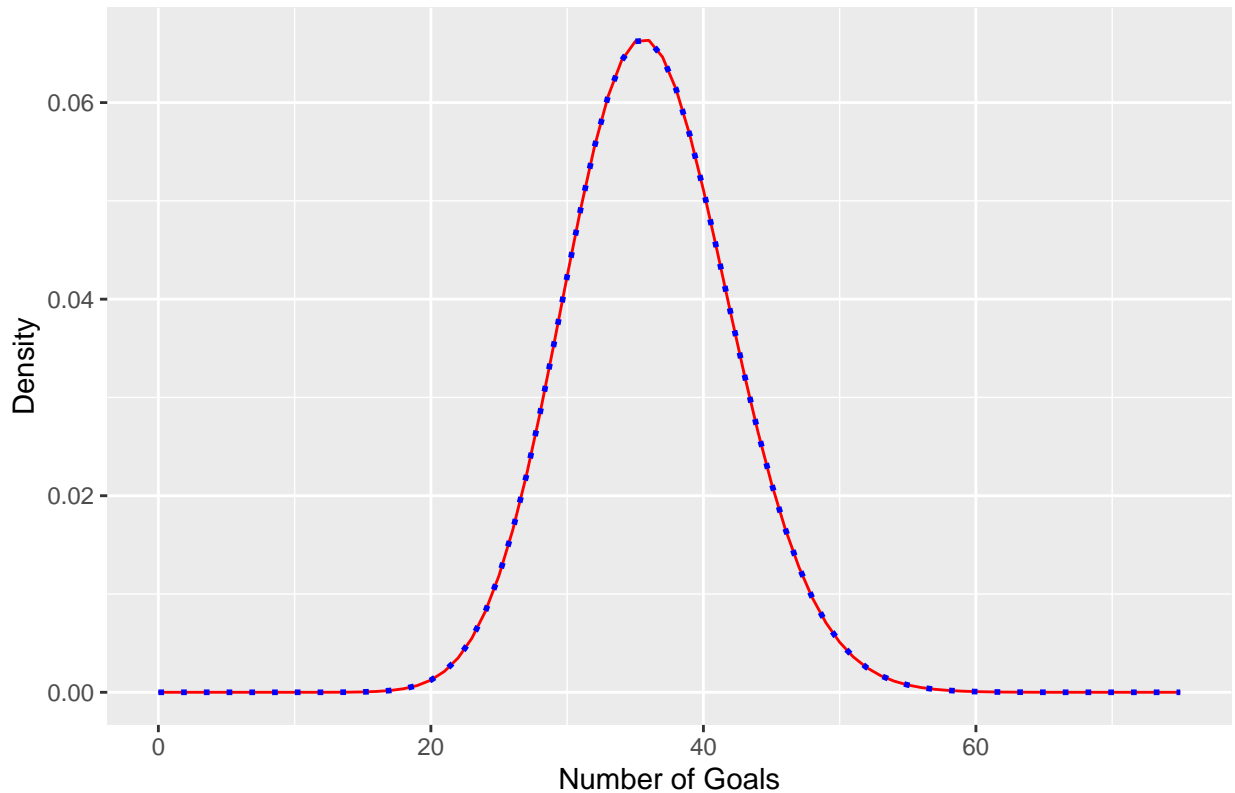
As calculated above, our P-value is 0.1156876. This means that over hypothetical replications, 11.6% of the time when  $H_0$  is true do we see a value of  $X$  that is as large or larger than 44 goals scored.

- c. Suppose Coach Babcock had said “He’s a really different player in the playoffs”. How would your answers to a and b change? Show your code for re-calculating the P-value.

If Coach Babcock had said “He’s a really different player in the playoffs”, our answer to a. and b. would change drastically. Unlike part a. and b., since Coach Babcock didn’t say he performs better in the playoffs, any unusual difference in play in the playoffs will be evidence to his statement. Thus, our alternative hypothesis will change to a two sided hypothesis:  $H_1 : \lambda_0 \neq \frac{206}{506}$ . However, there will be no change to the null hypothesis nor the sampling distribution of  $X$  under the null. As seen below, our sampling distribution is not perfectly symmetric, thus we can’t simply double our one-sided P-value, instead we will automate the two-sided P-value calculation using R.

```
ggplot(data.frame(x=c(0:75)), mapping = aes(x = x)) +
  geom_line(aes(y=dpois(x, 89 * 205/506)), color="red") +
  geom_line(aes(y=dpois(x, 89 * 205/506)), color="blue", linetype="dotted", linewidth = 1) +
  labs(title="Sampling Distribution of X under the Null Hypothesis",
       x = "Number of Goals",
       y = "Density")
```

## Sampling Distribution of X under the Null Hypothesis



```
probs <- dpois(x = 0:1000, lambda = 89 * 206/506)
twosidep <- sum(probs[probs <= dpois(x = 44, lambda = 89 * 206/506)])
```

Under our new alternative hypothesis the two-sided P-value calculated above is 0.2114414. Thus, under hypothetical replications, 21.1% of the time when  $H_0$  is true do we see values as or more unusual than observing 44 goals. This event is not very rare in the grand scheme of things.

It is important to note that since  $X$  can be thought of as the sum of 89 independent Poisson random variables, the Central Limit Theorem can be used to approximate the sampling distribution of  $X$  and thus P-values as well. This is seen by the near symmetric sampling distribution of  $X$ . However, if we were to double our one-sided P-value from part b. we'd notice that it is slightly higher/more conservative than the one calculated above, further showing the the distribution is not perfectly symmetric.

2. (Sign test) Suppose  $Y_1, Y_2, \dots, Y_{19}$  are independent random variables drawn from some distribution, and we are interested in the parameter  $\pi_0 = P(Y < 0)$  More precisely, we want to test  $H_0 : \pi_0 = 0.4$  against  $H_1 : \pi_0 < 0.4$  at the 5% level.

A reasonable test statistic would be  $X$ , the number of negative observations in the sample. That is,

$$X = X_1 + X_2 + \dots + X_{19}$$

where

$$X_i = \begin{cases} 1 & Y_i < 0 \\ 0 & \text{otherwise} \end{cases}$$

- a. What is the sampling distribution of  $X$  if the null hypothesis is true? Give a brief justification for your answer.

Since  $X_i$  can only take the values 1 if  $Y < 0$  and 0 otherwise, it is apparent that the  $X_i$ 's are independent and identically distributed Bernoulli random variables. Furthermore, since  $X$  is the sum of 19 Binomial random variables, it follows from Theorem 16.2 that  $X \sim \text{Binom}(n_1 + n_2 + \dots + n_{19}, \pi_0)$ . With this said, under  $H_0$  the value of  $\pi_0$  is 0.4. Therefore the sampling distribution of  $X$  if the null hypothesis is true is  $X \sim \text{Binom}(n = 19, \pi_0 = 0.4)$ .

- b. For what values of  $X$  would you reject the null hypothesis? Support your answer showing code/output as necessary.

First off, to find the rejection region, we need to find which tail  $H_1$  corresponds to. Since  $\pi_0$  represents the probability of getting a “success” and a success in our case is seeing a  $Y < 0$ , it follows that when  $\pi_0$  gets smaller the amount of  $Y$ 's  $< 0$  also gets smaller, thus we are looking at the lower tail. Furthermore, since  $\alpha = 0.05$  and we are looking at the lower tail, the critical value/cutoff point will be the number of successes such that 5% of the distribution is to the left of that number, this can be found using `qbinom`.

```
crit_val <- qbinom(p = 0.05, size = 19, prob = 0.4)
```

As computed above, our critical value from `qbinom` is 4. However, since the binomial distribution is a discrete distribution, the value given actually has a density of 0.06961371 to the left of it. With that said our rejection region is thus  $[0, 3]$ . Any values in this range will lead us to reject the null hypothesis.

- c. What is the Type I error probability of your test? Support your answer showing code/output as necessary.

Since a Type I error corresponds to rejecting the null hypothesis when the null hypothesis is true, in order to find the type 1 error rate for this example, we must find the probability of landing in the rejection region. This corresponds to computing the following probability statement:  $P(X \leq 3)$ . We can calculate this using `pbinom`.

```
type1 <- pbinom(q = 3, size = 19, prob = 0.4)
```

As computed above, the type 1 error rate/probability of seeing a type 1 error in the test is 0.0229593.

- d. Calculate the Type II error probability when  $\pi_0 = 0.2$ . Support your answer showing code/output as necessary.

Since a Type II error occurs when we fail to reject the null hypothesis when the alternative hypothesis is true, in order to find the Type II error rate for this example, we must find the probability of landing outside the rejection region, under the sampling distribution of  $H_1$ . In this case,  $\pi_0 = 0.2$ , and thus the sampling distribution of  $X$  under  $H_1$  is  $X \sim \text{Binom}(n = 19, \pi_0 = 0.2)$ . Finding  $P(\text{Type II error})$ , is the same as computing  $P(X > 3)$  under  $H_1$ . Again, this can be calculated using `pbinom`.

```
type2 <- pbinom(q = 3, size = 19, prob = 0.2, lower.tail = F)
```

As calculated above, when  $\pi_0 = 0.2$ , the Type II error probability is 0.5449113.

3. (Sign test again) Suppose in problem 2, you know that  $Y$  is normally distributed with mean  $\mu_0$  and (known) standard deviation  $\sigma_0 = 1$ .

- a. Re-state the null and alternative hypothesis from problem 2 in terms of claims about  $\mu_0$ .

In Problem 2, as stated above, we wanted to test  $H_0 : \pi_0 = 0.4$  against  $H_1 : \pi_0 < 0.4$  at the 5% level. However, now we know that  $Y$  is normally distributed with mean  $\mu_0$  and known standard deviation  $\sigma_0 = 1$ . In order to find what the mean value of  $Y$  should be, we must understand what  $\pi_0 = 0.4$  was telling us in the first place. In the context of the last problem,  $\pi_0 = 0.4$  was the success probability in which a success was seeing a  $Y < 0$ . Thus, in order to see what  $\mu_0$  should be in the null hypothesis, we need to find the mean value such that there is a probability of 0.4 of seeing a negative number. This can be done by creating a function including a `pnorm` statement and applying `uniroot` to find the mean value that satisfies the given condition.

```
mean_value <- function(x) {
  pnorm(q = 0, mean = x, sd = 1) - 0.4
}
```

```
}
null_mu <- uniroot(f = mean_value, lower = 0, upper = 1)$root
```

Since our alternative hypothesis in problem 2 corresponded to a smaller chance of seeing a value of  $Y < 0$ , in order to capture that same aspect of  $H_1$  in the case of testing  $\mu_0$ , we must look at values of  $\mu_0$  that are **greater** than the null, since this corresponds to smaller chances of seeing a negative number.

Our null and alternative hypotheses are thus:

$$H_0 : \mu_0 = 0.253 \quad H_1 : \mu_0 > 0.253$$

- b. We are still interested in performing the test at the 5% level, but we will now use the sample mean,  $\bar{Y}$ , as our statistic. What is the sampling distribution of  $\bar{Y}$  if the null hypothesis is true? Give a brief justification for your answer.

Unlike the last problem, we will now use  $\bar{Y}$ , as our statistic. Furthermore, since we are using  $\bar{Y}$  it follows from Slide 7 of the “Summary of Confidence Intervals” lecture notes that when  $Y \sim \text{Norm}(\mu_0, \sigma_0)$ , the sampling distribution of  $\bar{Y}$  is  $\bar{Y} \approx \text{Norm}(\mu_0, \frac{\sigma_0}{\sqrt{n}})$  for large  $n$ . If  $H_0$  is true and assuming  $n$  is large enough, the sampling distribution of  $\bar{Y}$  is  $\bar{Y} \sim \text{Norm}(\mu_0 = 0.253, \sigma_0 = \frac{1}{\sqrt{19}})$ .

- c. For what values of  $\bar{Y}$  will you reject the null hypothesis? Support your answer showing code/output as necessary.

First off, to find the rejection region, we need to find which tail  $H_1$  corresponds to. Since  $\mu_0$  represents the center of the normal distribution and larger  $\mu_0$  values correspond to moving farther to the right, we are thus looking at the upper tail. Furthermore, since  $\alpha = 0.05$  and we are looking at the upper tail, the critical value/cutoff point will be the number such that 5% of the distribution is to the right of that number, this can be found using `qnorm`.

```
crit_val2 <- qnorm(p = 0.95, mean = null_mu, sd = 1/sqrt(19))
```

Thus, we will reject the null hypothesis for any values greater than or equal to 0.6307008.

- d. What is the Type I error probability of your test? Briefly explain why.

Since the sampling distribution of  $\bar{Y}$  is continuous it follows that the probability of a Type I error is exactly the significance level, thus  $P(\text{Type I Error}) = 0.05$ .

To illustrate this we will use `pnorm` on our critical value calculated in part c.

```
type1cont <- pnorm(q = crit_val2, mean = null_mu, sd = 1/sqrt(19), lower.tail = F)
```

As computed above,  $P(\text{Type I Error}) = 0.05$ .

- e. Calculate the Type II error probability when  $\pi_0 = 0.2$ . Support your answer showing code/output as necessary.

Just as we did in part a., we need to find the value of  $\mu_0$  that corresponds to  $\pi_0 = 0.2$ . Again, we will make a function using `pnorm` and use `uniroot` to solve it.

```
mean_value_alt <- function(x) {
  pnorm(q = 0, mean = x, sd = 1) - 0.2
}

alt_mu <- uniroot(f = mean_value_alt, lower = 0, upper = 1)$root
```

Doing this, we see that  $\pi_0 = 0.2$  corresponds to seeing a  $\mu_0$  take on a value of 0.8416441.

Since a Type II error occurs when we fail to reject the null hypothesis when the alternative hypothesis is true, in order to find the Type II error rate for this example, we must find the probability of landing outside

the rejection region, under the sampling distribution of  $H_1$ . In this case,  $\mu_0 = 0.842$ , and thus the sampling distribution of  $\bar{Y}$  under  $H_1$  is  $\bar{Y} \sim \text{Norm}(\mu_0 = 0.842, \sigma_0 = \frac{1}{\sqrt{19}})$ . Finding  $P(\text{Type II error})$ , is the same as computing  $P(\bar{Y} < 0.6307008)$  under  $H_1$ . Again, this can be calculated using `pnorm`.

```
type2cont <- pnorm(q = crit_val2, mean = alt_mu, sd = 1/sqrt(19))
```

As calculated above, when  $\mu_0 = 0.842$ , the Type II error probability is 0.1789221.