

## Discretionary Note

Anish Krishna Lakkapragada

**IF YOU USE THIS FILE TO CHEAT, YOU ARE NOT ONLY STUPID BUT YOU ARE CHEATING YOURSELF OUT OF THE ABILITY TO FALL IN LOVE WITH MATH.** Furthermore, I am not smarter than you and my solutions did not always get a perfect score.

**CONTENT STARTS ON NEXT PAGE.**

To access the general instructions for this repository head [here](#).

# STATS 242 HW 4

February 11, 2025

---

Number of late days: 0; Collaborators: Derek Gao, Alex Wa

1.

- (a) A pivotal statistic is one where the test statistic null distribution does not change regardless of the parameters of the sample data. Because under  $H_0$ , **(1)**  $S_i \sim \text{Bern}(\frac{1}{2})$  as any PDF  $f$  (regardless of its parameters) is symmetric around 0 (i.e.  $P(X_i > 0) = 50\%$ ) and **(2)** any given  $X_i$  is equally likely to have a rank of 1 to  $n$  (i.e.  $R_i \sim \text{Unif}(0, 1)$ ), we can see that  $W = \sum_{i=1}^n S_i R_i$  does not depend on any parameters of  $X_i \implies W$  is pivotal under  $H_0$ .

If the  $X_i$ 's tended to take positive values, then the  $S_i$ 's would be more likely to be one than zero and so  $W = \sum_{i=1}^n S_i R_i$  would count more of the (strictly positive)  $R_i$ 's so  $W$  would be larger. To test against a one-sided alternative  $H_1$  that the  $X_i$ 's tended to take positive values, I would reject  $H_0$  for large values of  $W$ .

- (b) We can represent  $W = \sum_{k=1}^n k I_k$ , where  $I_k$  is one if observation  $i$  with rank  $R_i = k$  has  $S_i = 1$  and zero otherwise. In other words, this summation is essentially an enumeration over ranks  $1 \dots n$  that only sums the ranks of data points where  $X_i \geq 0 \implies S_i = 1$ . Note that under  $H_0$ ,  $\forall k \in [1, n]$ ,  $I_k \sim \text{Bern}(\frac{1}{2})$  as we are assuming  $f$  is symmetric around zero which means there is a 50% chance  $X_k \geq 0$ . Given this, we compute the expectation of  $W$  below<sup>1</sup>:

$$\mathbb{E}[W] = \mathbb{E}[\sum_{k=1}^n k I_k] = \sum_{k=1}^n \mathbb{E}[k I_k] = \sum_{k=1}^n k \mathbb{E}[I_k] = \frac{1}{2} \sum_{k=1}^n k = \frac{1}{2} \frac{n(n+1)}{2} = \frac{n(n+1)}{4}$$

We now compute the  $\text{Var}(W)$ . Note that because each  $X_i$  is independent, each  $I_k$  is independent and so  $\text{Var}(W) = \text{Var}(\sum_{k=1}^n k I_k) = \sum_{k=1}^n \text{Var}(k I_k)$ . We compute the variance of  $W$  below<sup>2</sup>:

$$\begin{aligned} \text{Var}(W) &= \sum_{k=1}^n \text{Var}(k I_k) = \sum_{k=1}^n k^2 \text{Var}(I_k) = \text{Var}(I_k) \sum_{k=1}^n k^2 = \frac{1}{2} \left(1 - \frac{1}{2}\right) \sum_{k=1}^n k^2 \\ &= \frac{1}{4} \sum_{k=1}^n k^2 = \frac{1}{4} \frac{n(n+1)(2n+1)}{6} = \frac{n(n+1)(2n+1)}{24} \end{aligned}$$

---

<sup>1</sup>We use the fact that  $\sum_{k=1}^n k = \frac{n(n+1)}{2}$ .

<sup>2</sup>We use the fact that  $\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$ .

Let us assume that under large  $n$ ,  $W$  can be approximated by  $\mathcal{N}(\frac{n(n+1)}{4}, \frac{n(n+1)(2n+1)}{24})$ , which we will refer to by  $W_n$ . As stated in part (a), if  $X_i$ 's tended to take positive values, we would reject  $H_0$  for large values of  $W$ . This means that to perform this test at  $\alpha$  significance level, we would first find the upper- $\alpha$  point  $z^\alpha$  of  $W^3$  and then if test statistic  $W$  for  $X_1, \dots, X_n$  is greater than  $z^\alpha$ , we reject  $H_0$ .

2.

- (a) Given  $|X_1|, \dots, |X_n|$ , we have no information on the signed values  $X_1, \dots, X_n$ . Given  $|X_i|$ , there are two possible values of  $X_i$  (i.e.  $|X_i|$  or  $-|X_i|$ ). Thus, given  $|X_1|, \dots, |X_n|$ , there are  $2^n$  possible values of the set  $X_1, \dots, X_n$ . This means that the distribution of  $T$  conditional on  $|X_1|, \dots, |X_n|$  can take (at most)  $2^n$  unique values<sup>4</sup> as there are  $2^n$  possible unique configurations of  $X_1, \dots, X_n$ . The probability of any of these unique values of  $T$  is given by  $\frac{k}{2^n}$  where  $k$  is the number of times this value occurs as the evaluation of  $T$  across all configurations of  $X_1, \dots, X_n$ . The  $\frac{1}{2^n}$  term reflects the fact that each of these configurations are equally likely<sup>5</sup>.
- (b) To conduct a level- $\alpha$  test that rejects  $H_0$  for large values of  $T$ , we first have to find the null distribution of  $T$ . We do this with computer simulation. Because we are not given the PDF  $f$ , we cannot just repeatedly sample values from this distribution. However, we are given a set  $X_1, \dots, X_n$  that is realized from this unknown distribution. Under  $H_0$ , this set of data  $X_1, \dots, X_n$  is just as likely as any set of  $\pm X_1, \dots, \pm X_n$ . Thus, we can compute  $T$  for all sign permutations of  $X_1, \dots, X_n$ . From this set of values of  $T$ , we have an approximation of the null distribution of  $T$  and thus can take the top  $(100 - \alpha)$ th percentile of  $T$  as the upper- $\alpha$  point of  $T$ . If  $T(X_1, \dots, X_n) >$  this upper- $\alpha$  point of  $T$ , then we will reject  $H_0$ . If  $Y_1, \dots, Y_n$  and  $Z_1, \dots, Z_n$  are each  $n$  IID data points and each  $X_i$  is given by  $Y_i - Z_i$ , then the  $H_0$  that  $f$  is symmetric around zero  $\implies$  each  $X_i$  follows the same distribution as  $-X_i = Z_i - Y_i \implies X_i$  has the same distribution regardless of if it is computed on  $(Y_i, Z_i)$  or  $(Z_i, Y_i) \implies (Y_i, Z_i)$  and  $(Z_i, Y_i)$  have the same (bivariate) distribution.

3.

Because both the null and alternative distributions, given by  $f_0(x)$  and  $f_1(x)$ , are fully specified (i.e. no unknown parameters), they are both simple hypotheses and so we can apply the Neyman-Pearson Lemma. The Neyman-Pearson Lemma tells us that if we can find a  $c$  such that the Type I error probability is equal to  $\alpha = 0.10$ , the likelihood ratio test<sup>6</sup> is guaranteed to be the test with the highest power. Thus, we first solve for  $c$ , which is the upper- $\alpha$  point of the likelihood ratio test statistic  $L(x) = \frac{f_1(x)}{f_0(x)} = 2x$  where  $x \in [0, 1]$  and  $L(x) \in [0, 2]$ :

<sup>3</sup>More formally,  $z^\alpha$  is given by  $\int_{z^\alpha}^{\infty} f_{W_n}(w)dw = \alpha$ , where  $f_{W_n}$  is the PDF of  $W_n$ .

<sup>4</sup>Expressed differently, the  $2^n$  values  $T$  can take are  $T(\pm X_1, \pm X_2, \dots, \pm X_n)$ .

<sup>5</sup>This is for two reasons: **(1)** all  $X_i$  are independent and **(2)** under  $H_0$ ,  $f$  is symmetric around zero and so  $X_i$  is equally likely to be positive or negative.

<sup>6</sup>using  $c$  to define the rejection region

$$\begin{aligned}
\mathbb{P}[\text{Type I Error}] &= \mathbb{P}_{H_0}[\text{reject } H_0] = \mathbb{P}_{H_0}[L(x) > c] = \alpha = 0.10 \\
\mathbb{P}_{H_0}[L(x) > c] &= \mathbb{P}_{H_0}[2x > c] = \mathbb{P}_{H_0}[x > \frac{c}{2}] = \int_{0.5c}^1 f_0(x)dx = 1 - 0.5c = 0.10 \\
c &= 1.8
\end{aligned}$$

Thus, we will reject any sample  $x$  when  $L(x) > c$ . Given this threshold, we can compute the power of the test, which is given by  $\mathbb{P}_{H_1}[\text{reject } H_0] = \mathbb{P}_{H_1}[L(X) > c]$ :

$$\begin{aligned}
\mathbb{P}_{H_1}[L(X) > c] &= \mathbb{P}_{H_1}[2X > c] = \mathbb{P}_{H_1}[X > \frac{c}{2}] = \int_{0.5c}^1 f_1(x)dx \\
&= \int_{0.5c}^1 2x dx = x^2 \Big|_{0.5c}^1 = 1 - 0.25c^2 = 1 - 1.8^2(0.25) = 0.19
\end{aligned}$$

Thus the maximum power of a test with these hypotheses at the significance level  $\alpha = 0.1$  significance level is 19%.

4.

- a) Because both  $\sigma_0^2$  and  $\sigma_1^2$  are known,  $H_0$  and  $H_1$  are simple hypotheses. So the Neyman-Pearson Lemma applies in this scenario and guarantees that the likelihood ratio test is the most powerful test. Let us define  $f_0(x)$  be the PDF of  $X \sim \mathcal{N}(0, \sigma_0^2)$  under  $H_0$  and  $f_1(x)$  be the PDF of  $X \sim \mathcal{N}(0, \sigma_1^2)$  under  $H_1$ . Furthermore, let vector  $\mathbf{x} = (X_1, \dots, X_n)$ . Then the likelihood ratio test statistic on  $X_1, \dots, X_n$  is given by  $L(\mathbf{x}) = \frac{f_1(\mathbf{x})}{f_0(\mathbf{x})}$ . We compute this below:

$$\begin{aligned}
f_0(\mathbf{x}) &= \prod_{i=1}^n f_0(x_i) = \prod_{i=1}^n \frac{\exp(\frac{-x_i^2}{2\sigma_0^2})}{\sqrt{2\pi}\sigma_0} = (\frac{1}{\sqrt{2\pi}\sigma_0})^n \exp(\frac{-1}{2\sigma_0^2}[x_1^2 + \dots + x_n^2]) \\
f_1(\mathbf{x}) &= \prod_{i=1}^n f_1(x_i) = \prod_{i=1}^n \frac{\exp(\frac{-x_i^2}{2\sigma_1^2})}{\sqrt{2\pi}\sigma_1} = (\frac{1}{\sqrt{2\pi}\sigma_1})^n \exp(\frac{-1}{2\sigma_1^2}[x_1^2 + \dots + x_n^2]) \\
L(\mathbf{x}) &= \frac{f_1(\mathbf{x})}{f_0(\mathbf{x})} = (\frac{\sigma_1}{\sigma_0})^n \exp([\frac{1}{2\sigma_0^2} - \frac{1}{2\sigma_1^2}][x_1^2 + \dots + x_n^2]) \\
&= (\frac{\sigma_1}{\sigma_0})^n \exp(\frac{2\sigma_1^2 - 2\sigma_0^2}{4\sigma_0^2\sigma_1^2}[x_1^2 + \dots + x_n^2]) = (\frac{\sigma_1}{\sigma_0})^n \exp(\frac{1}{2} \frac{\sigma_1^2 - \sigma_0^2}{\sigma_0^2\sigma_1^2}[x_1^2 + \dots + x_n^2])
\end{aligned}$$

We can observe that for  $\sigma_1^2 > \sigma_0^2 \implies \frac{1}{2\sigma_0^2} - \frac{1}{2\sigma_1^2} > 0$ , the test statistic  $L(\mathbf{x})$  is thus an increasing function of  $x_1^2 + \dots + x_n^2$ . Under  $H_0$ , each  $x_i \sim \mathcal{N}(0, \sigma_0^2)$ , and so to standardize  $x_1^2 + \dots + x_n^2$ , we can say that  $L(\mathbf{X})$  is an increasing function of  $\frac{1}{\sigma_0^2}(x_1^2 + \dots + x_n^2)$ , which we can define as  $X_n^2$ . Note that because each  $\frac{X_i}{\sigma_0} \sim \mathcal{N}(0, 1)$ ,

$X_n^2 \sim \chi_n^2$ . Because  $L(\mathbf{X})$  is an increasing function of  $X_n^2$ , this means the rejection event  $L(\mathbf{x}) > \text{upper-}\alpha \text{ point of } L(\mathbf{X}) \text{ null distribution}$  is *equivalent* to the rejection event  $X_n^2 > \text{upper-}\alpha \text{ point of its distribution, } \chi_n^2$ . This upper- $\alpha$  point is given to us by  $\chi_n^2(\alpha)$ .

Given this, we can define a test statistic  $T(\mathbf{x})$ :

$$T(\mathbf{x}) = \frac{x_1^2 + \cdots + x_n^2}{\sigma_0^2}$$

and the rejection region  $\mathcal{R}$  for this test can be defined as:

$$\mathcal{R} = \{x : T(x) > \chi_n^2(\alpha)\}$$

- b) Under the alternative hypothesis  $H_1$ , each  $X_i \sim \mathcal{N}(0, \sigma_1^2)$ . This means that  $\frac{x_1^2 + \cdots + x_n^2}{\sigma_1^2}$  follows a  $\chi_n^2$  distribution, and so:

$$T(\mathbf{x}) = \frac{x_1^2 + \cdots + x_n^2}{\sigma_0^2} = \frac{x_1^2 + \cdots + x_n^2}{\sigma_1^2} \cdot \frac{\sigma_1^2}{\sigma_0^2} \sim \frac{\sigma_1^2}{\sigma_0^2} \chi_n^2$$

We now solve for the power of this test, which is given by  $\mathbb{P}_{H_1}[\text{reject } H_0]$ :

$$\begin{aligned} \mathbb{P}_{H_1}[\text{reject } H_0] &= \mathbb{P}_{H_1}[T(\mathbf{x}) > \chi_n^2(\alpha)] = \mathbb{P}_{H_1}\left[\frac{\sigma_1^2}{\sigma_0^2} \chi_n^2 > \chi_n^2(\alpha)\right] = \mathbb{P}_{H_1}\left[\chi_n^2 > \frac{\sigma_0^2}{\sigma_1^2} \chi_n^2(\alpha)\right] \\ &= 1 - F\left(\frac{\sigma_0^2}{\sigma_1^2} \chi_n^2(\alpha)\right) \end{aligned}$$

Keeping  $\sigma_0^2$  and  $\alpha$  fixed, we can see that as  $\sigma_1^2 \rightarrow \infty$ ,  $\frac{\sigma_0^2}{\sigma_1^2} \chi_n^2(\alpha)$  goes closer to zero and so  $F\left(\frac{\sigma_0^2}{\sigma_1^2} \chi_n^2(\alpha)\right)$  also goes closer to zero which means the power given by  $1 - F\left(\frac{\sigma_0^2}{\sigma_1^2} \chi_n^2(\alpha)\right)$  approaches one.