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 $W = \Sigma_{i=1}^n S_i R_i$ 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 $W = \Sigma_{i=1}^n S_i R_i$ would count more of the (stricly 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 $W = \Sigma_{k=1}^n k I_k$, 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 $1 \ldots n$ 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[1]:

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

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

$$\text{Var}(W) = \Sigma_{k=1}^n \text{Var}(k I_k) = \Sigma_{k=1}^n k^2 \text{Var}(I_k) = \text{Var}(I_k)\Sigma_{k=1}^n k^2 = \frac{1}{2}(1 - \frac{1}{2})\Sigma_{k=1}^n k^2$$
$$= \frac{1}{4}\Sigma_{k=1}^n k^2 = \frac{1}{4}\frac{n(n+1)(2n+1)}{6} = \frac{n(n+1)(2n+1)}{24}$$

[1] We use the fact that $\Sigma_{k=1}^n k = \frac{n(n+1)}{2}$.

[2] We use the fact that $\Sigma_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$.

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 α significance level, we would first find the upper-α point z^α of W[3] and then if test statistic W for $X_1, \ldots, X_n$ is greater than z^α, we reject H_0.

2.

(a) Given $|X_1|, \ldots, |X_n|$, we have no information on the signed values $X_1, \ldots, X_n$. Given $|X_i|$, there are two possible values of X_i (i.e. $|X_i|$ or $-|X_i|$). Thus, given $|X_1|, \ldots, |X_n|$, there are 2^n possible values of the set $X_1, \ldots, X_n$. This means that the distribution of T conditional on $|X_1|, \ldots, |X_n|$ can take (at most) 2^n unique values[4] as there are 2^n possible unique configurations of $X_1, \ldots, X_n$. 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 $X_1, \ldots, X_n$. The $\frac{1}{2^n}$ term reflects the fact that each of these configurations are equally likely[5].

(b) To conduct a level-α 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 $X_1, \ldots, X_n$ that is realized from this unknown distribution. Under H_0, this set of data $X_1, \ldots, X_n$ is just as likely as any set of $\pm X_1, \ldots, \pm X_n$. Thus, we can compute T for all sign permutations of $X_1, \ldots, X_n$. 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-α point of T. If $T(X_1, \ldots, X_n) >$ this upper-α point of T, then we will reject H_0.

If $Y_1, \ldots, Y_n$ and $Z_1, \ldots, Z_n$ 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[6] is guaranteed to be the test with the highest power. Thus, we first solve for c, which is the upper-α point of the likelihood ratio test statistic $L(x) = \frac{f_1(x)}{f_0(x)} = 2x$ where $x \in [0, 1]$ and $L(x) \in [0, 2]$:

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

[4] Expressed differently, the 2^n values T can take are $T(\pm X_1, \pm X_2, \ldots, \pm X_n)$.

[5] 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.

[6] using c to define the rejection region

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

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

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

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 σ_0^2 and σ_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 $\mathbf{x} = (X_1, \ldots, X_n)$. Then the likelihood ratio test statistic on $X_1, \ldots, X_n$ is given by $L(\mathbf{x}) = \frac{f_1(\mathbf{x})}{f_0(\mathbf{x})}$. We compute this below:

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

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 $x_1^2 + \cdots + x_n^2$. Under H_0, each $x_i \sim \mathcal{N}(0, \sigma_0^2)$, and so to standardize $x_1^2 + \cdots + x_n^2$, we can say that $L(\mathbf{X})$ is an increasing function of $\frac{1}{\sigma_0^2}(x_1^2 + \cdots + x_n^2)$, 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 $L(\mathbf{x}) >$ upper-α point of $L(\mathbf{X})$ null distribution is *equivalent* to the rejection event $X_n^2 >$ upper-α point of its distribution, χ_n^2. This upper-α 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 text 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 χ_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 $\mathbb{P}_{H_1}[\text{reject } H_0]$:

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

Keeping σ_0^2 and α fixed, we can see that as $\sigma_1^2 \to \infty$, $\frac{\sigma_0^2}{\sigma_1^2}\chi_n^2(\alpha)$ goes closer to zero and so $F(\frac{\sigma_0^2}{\sigma_1^2}\chi_n^2(\alpha))$ also goes closer to zero which means the power given by $1-F(\frac{\sigma_0^2}{\sigma_1^2}\chi_n^2(\alpha))$ approaches one.