STATS 200: Homework 3

Due Wednesday, October 19, at 5PM

1. **The t_1 distribution.**
(a) Let $T \sim t_1$ (the t distribution with 1 degree of freedom). Explain why T has the same distribution as $\frac{X}{|Y|}$ where $X, Y \overset{IID}{\sim} \mathcal{N}(0,1)$, and hence why T also has the same distribution as $\frac{X}{Y}$.

(The distribution of $\frac{X}{Y}$ when $X, Y \overset{IID}{\sim} \mathcal{N}(0,1)$ is also called the Cauchy distribution. An exercise using the change-of-variables formula shows that this has PDF

$$f(x) = \frac{1}{\pi}\frac{1}{x^2+1}.$$

You may use this result without proof in part (b).)

(b) t_1 is an example of an extremely "heavy-tailed" distribution: For $T \sim t_1$, show that $\mathbb{E}[|T|] = \infty$ and $\mathbb{E}[T^2] = \infty$. If $T_1, \ldots, T_n \overset{IID}{\sim} t_1$, explain why the Law of Large Numbers and the Central Limit Theorem do not apply to the sample mean $\frac{1}{n}(T_1 + \ldots + T_n)$.

2. **The t_n distribution for large n.**
For this question, you may use a result called Slutsky's lemma: If sequences of random variables $\{X_n\}_{n=1}^{\infty}$ and $\{Y_n\}_{n=1}^{\infty}$ satisfy $X_n \to c$ in probability for a constant $c \in \mathbb{R}$ and $Y_n \to Y$ in distribution for a random variable Y, then $X_n Y_n \to cY$ in distribution.

(a) Let $U_n \sim \chi^2_n$. Show that $1/\sqrt{\frac{1}{n}U_n} \to 1$ in probability as $n \to \infty$. (Hint: Apply the Law of Large Numbers and the Continuous Mapping Theorem.)

(b) Using Slutsky's lemma, show that if $T_n \sim t_n$ for each $n = 1, 2, 3, \ldots$, then as $n \to \infty$, $T_n \to Z$ in distribution where $Z \sim \mathcal{N}(0,1)$. (This formalizes the statement that "the t_n distribution approaches the standard normal distribution as n gets large".)

(c) Explain heuristically, in a few sentences, why use of the t-test is approximately valid by the Central Limit Theorem if n is large, even if $X_1, \ldots, X_n$ are not normally distributed. (You may use the fact that when n is large, the ratio of the sample variance S^2 to the true data variance σ^2 is very close to 1 with high probability.)

3. **Comparing binomial proportions.** The internet company Oogl would like to understand whether visitors to a website are more likely to click on an advertisement at the top

of the page than one on the side of the page. They conduct an "AB test" in which they show n visitors (group A) a version of the website with the advertisement at the top, and m visitors (group B) a version of the website with the (same) advertisement at the side. They record how many visitors in each group clicked on the advertisement.

(a) Formulate this problem as a hypothesis test. (You may assume that visitors in group A independently click on the ad with probability p_A and visitors in group B independently click on the ad with probability p_B, where both p_A and p_B are unknown probabilities in $(0,1)$.) What are the null and alternative hypotheses? Are they simple or composite?

(b) Let $\hat{p}_A$ be the fraction of visitors in group A who clicked on the ad, and similarly for $\hat{p}_B$. A reasonable intuition is to reject H_0 when $\hat{p}_A - \hat{p}_B$ is large. What is the variance of $\hat{p}_A - \hat{p}_B$? Is this the same for all data distributions in H_0?

(c) Describe a way to estimate the variance of $\hat{p}_A - \hat{p}_B$ using the available data, assuming H_0 is true—call this estimate $\hat{V}$. Explain heuristically why, when n and m are both large, the test statistic

$$T = \frac{\hat{p}_A - \hat{p}_B}{\sqrt{\hat{V}}}$$

is approximately distributed as $\mathcal{N}(0,1)$ under any data distribution in H_0. (You may assume that when n and m are both large, the ratio of $\hat{V}$ to the true variance of $\hat{p}_A - \hat{p}_B$ that you derived in part (b) is very close to 1 with high probability.) Explain how to use this observation to perform an approximate level-α test of H_0 versus H_1.

4. **Sign test.** Consider data $X_1, \ldots, X_n \overset{IID}{\sim} f$ for some unknown probability density function f, and the testing problem

$$\begin{aligned} H_0 &: f \text{ has median } 0 \\ H_1 &: f \text{ has median } \mu \text{ for some } \mu > 0 \end{aligned}$$

(a) Explain why the Wilcoxon signed rank statistic does not have the same sampling distribution under every $P \in H_0$. Draw a picture of the graph of a density function f with median 0, such that the Wilcoxon signed rank statistic would tend to take larger values under f than under any density function g that is symmetric about 0.

(b) Consider the **sign statistic** S, defined as the number of values in $X_1, \ldots, X_n$ that are greater than 0. Explain why S has the same sampling distribution under every $P \in H_0$. How would you conduct a level-α test of H_0 versus H_1 using the test statistic S? (Describe explicitly the rejection threshold; you may assume that for $X \sim \text{Binomial}(n, \frac{1}{2})$, there exists an integer k such that $\mathbb{P}[X \geq k]$ is exactly α.)

(c) When n is large, explain why we may reject H_0 when $S > \frac{n}{2} + \sqrt{\frac{n}{4}}z(\alpha)$ where $z(\alpha)$ is the upper α point of $\mathcal{N}(0,1)$, instead of using the rejection threshold you derived in part (b).

5. **Power comparisons.** Consider the problem of testing

$$H_0 : X_1, \ldots, X_n \overset{IID}{\sim} \mathcal{N}(0, 1)$$
$$H_1 : X_1, \ldots, X_n \overset{IID}{\sim} \mathcal{N}(\mu, 1)$$

at significance level $\alpha = 0.05$, where $\mu > 0$. We've seen four tests that may be applied to this problem, summarized below:

- Likelihood ratio test: Reject H_0 when $\bar{X} > \frac{1}{\sqrt{n}} z(0.05)$.
- t-test: Reject H_0 when $T := \sqrt{n}\bar{X}/S > t_{n-1}(0.05)$, where $S^2 = \frac{1}{n-1}\sum_i (X_i - \bar{X})^2$.
- Wilcoxon signed rank test: Reject H_0 when $W_+ > \frac{n(n+1)}{4} + \sqrt{\frac{n(n+1)(2n+1)}{24}} z(0.05)$, where W_+ is the Wilcoxon signed rank statistic.
- Sign test (from Problem 4 above): Reject H_0 when $S > \frac{n}{2} + \sqrt{\frac{n}{4}} z(0.05)$, where S is the number of positive values in $X_1, \ldots, X_n$.

(For the Wilcoxon and sign test statistics, we are using the normal approximations for their null distributions.) These tests are successively more robust to violations of the $\mathcal{N}(0, 1)$ distributional assumption imposed by H_0.

(a) For $n = 100$, verify numerically that these tests have significance level close to α, in the following way: Perform 10,000 simulations. In each simulation, draw a sample of 100 observations from $\mathcal{N}(0, 1)$, compute the above four test statistics $\bar{X}$, T, W_+, and S on this sample, and record whether each test accepts or rejects H_0. Report the fraction of simulations for which each test rejected H_0, and confirm that these fractions are close to 0.05.

For those of you doing this in R, the following commands may be helpful:

```
qnorm(0.95)
qt(0.95,df=99)
```

give the values $z(0.05)$ and $t_{99}(0.05)$, respectively (i.e. the 0.95 quantiles of the $\mathcal{N}(0, 1)$ and t_{99} distributions). For a numeric data vector X, you may use the built-in functions

```
t.test(X)$statistic
wilcox.test(X)$statistic
```

to compute the values of the t-statistic and the Wilcoxon signed-rank statistic. (Computing $\bar{X}$ and S should be easy using the functions reviewed in the last two homeworks.) To check whether a test accepts or rejects, we may use an if-else statement. For example, the following records whether the Wilcoxon test rejects across each of 10,000 simulations:

```
n = 100
output.W = numeric(10000)
for (i in 1:10000) {
```

```
  X = rnorm(n, mean=0, sd=1)
  W = wilcox.test(X)$statistic
  if (W > n*(n+1)/4+sqrt(n*(n+1)*(2*n+1)/24)*qnorm(0.95)) {
    output.W[i] = 1
  } else {
    output.W[i] = 0
  }
}
```

(b) For $n = 100$, numerically compute the powers of these tests against the alternative H_1, for the values $\mu = 0.1$, 0.2, 0.3, and 0.4. Do this by performing 10,000 simulations as in part (a), except now drawing each sample of 100 observations from $\mathcal{N}(\mu, 1)$ instead of $\mathcal{N}(0, 1)$. (You should be able to re-use most of your code from part (a).) Report your computed powers either in a table or visually using a graph.

(c) How do the powers of the four tests compare, when testing against a normal alternative? Your friend says, "We should always use the testing procedure that makes the fewest distributional assumptions, because we never know in practice, for example, whether the variance is truly 1 or whether data is truly normal." Comment on this statement. Rice says, "It has been shown that even when the assumption of normality holds, the [Wilcoxon] signed rank test is nearly as powerful as the t test. The [signed rank test] is thus generally preferable, especially for small sample sizes." Do your simulated results support this conclusion?