Name: ____________________ NetID: [ ][ ][ ][ ][ ][ ][ ][ ][ ][ ][ ]

Name: ____________________ NetID: [ ][ ][ ][ ][ ][ ][ ][ ][ ][ ][ ]

Name: ____________________ NetID: [ ][ ][ ][ ][ ][ ][ ][ ][ ][ ][ ]

Name: ____________________ NetID: [ ][ ][ ][ ][ ][ ][ ][ ][ ][ ][ ]

Work in groups of at least 2 and at most 4.

Suppose $X_1, \dots, X_n \sim N(\mu, 1)$ where μ is some unknown fixed parameter. We will assume $\Theta = \mathbb{R}$; i.e. that the mean could be anything.

Our goal is to design and analyze a test for the null hypothesis: *the mean μ is less than or equal to zero.*

1. What is Θ_0 and Θ_0^c?

iWe want $\Theta_0 = \{x : x \leq 0\}$ and $\Theta_0^c = \{x : x > 0\}$.

Define $T(x_1, \dots, x_n) = n^{-1}(x_1 + \cdots + x_n)$. Our test is the following:

- if $T(X_1, \dots, X_n) > c$: reject the null hypothesis
- if $T(X_1, \dots, X_n) \leq c$: retain the null hypothesis

In other words, the rejection region is $\{(x_1, \dots, x_n) \in \mathbb{R}^n : T(x_1, \dots, x_n) > c\}$.

2. How can we choose c to make Type I errors unlikely? What about to make Type II errors unlikely?

If we make c really large, then we will reject the null hypothesis less frequently and avoid Type I errors. Alternately, if we make c really small, we will retain the null hypothesis less frequently, and avoid Type II errors.

3. Find the power function $\beta(\mu) = \mathbb{P}[T(X_1, \dots, X_n) > c | X_1, \dots, X_n \sim N(\mu, 1)]$. You can write your answer in terms of $\Phi(x) = \mathbb{P}[Z \le x | X \sim N(0,1)]$ and use the fact that if $Y \sim N(\mu_Y, \sigma_Y^2)$ and $Z \sim N(\mu_Z, \sigma_Z^2)$ are independent, then $Y + Z \sim N(\mu_Y + \mu_Z, \sigma_Y^2 + \sigma_Z^2)$.

$$
\begin{aligned}
\mathbb{P}[T(X_1, \dots, X_n) > c | X_1, \dots, X_n \sim N(\mu, 1)] &= \mathbb{P}[n^{-1}(X_1 + \dots + X_n) > c | X_1, \dots, X_n \sim N(\mu, 1)] \\
&= \mathbb{P}[n^{-1} X > c | X \sim N(n\mu, n)] \\
&= \mathbb{P}[n^{-1}(X - n\mu) > c | X \sim N(0, n)] \\
&= \mathbb{P}[n^{-1}(n^{1/2} X + n\mu) > c | X \sim N(0, 1)] \\
&= \mathbb{P}[n^{-1/2} X + \mu > c | X \sim N(0, 1)] \\
&= \mathbb{P}[X > n^{1/2}(c - \mu) | X \sim N(0, 1)] \\
&= 1 - \Phi(n^{1/2}(c - \mu)).
\end{aligned}
$$

4. What is the size of the test? Given $\alpha \in (0, 1)$, for what values of c does the test have level α?

The size is $\sup_{\mu \in \Theta_0} \beta(\mu) = \sup_{\mu \le 0} \beta(\theta) = \beta(0) = 1 - \Phi(n^{1/2} c)$.

The test is of level-α if $c \ge n^{-1/2} \Phi^{-1}(1 - \alpha)$.

5. For each α we will pick $c = c_\alpha$ so that our test is a size-α test. Suppose $n = 8$ and we sample data points $\{3, -1, 5, 0, 3, -4, -1, 2\}$. What is the corresponding p-value?

   What if we had the same sample mean, but $n = 100$? Before doing the computation, think about if the p-value should be smaller or larger.

   What if instead we had data $\{0, 2, -1, 1, -2, -3, 4, -1\}$?

   In Wolfram Alpha you can evaluate $\Phi(1.24)$ using the query “probabilities for normal 1.24”.

For each α, we have $c = n^{-1/2} \Phi^{-1}(1 - \alpha)$.

We have $\bar{X}_8 = /8$, so we will reject when $\bar{X}_8 > n^{-1/2} \Phi^{-1}(1 - \alpha)$, or equivalently, when $\alpha < 1 - \Phi(n^{1/2} \bar{X}_8) \approx 0.0067$.

In this last case, p is 1.06×10^{-18}. This makes sense, because if the mean of 100 data points was $7/8$, we are more sure the true mean is > 0 than if we just have 8 data points.

In the final case, the sample mean is 0 so the p value is 1/2.

6. Suppose instead, for each α, we pick c twice what we used in the previous part. How do the p-values change?

We now reject when $\bar{X}_8 > 2n^{-1/2} \Phi^{-1}(1 - \alpha)$, or equivalently, when $\alpha < 1 - \Phi(n^{1/2} \bar{X}_8 / 2) \approx 0.108$.