Part 1

1

Briefly describe in your own words the purpose of a hypothesis test and how you use your data to conclude whether you will or will not reject the null hypothesis.
Answer:

The big thing going on with these tests is we want to gauge the likhelihood of our data given some null hypothesis being true. Say we assume some average weight of a breed of dog, then we collect a lot of data and get a sample average that seems very different. Hypothesis tests give us a way to actually test that difference, to see how big of one it actually is. Is it likely that we got the sample average we did due to just random chance? What's the probability we actually got that value or a value even more extreme than it? That's what is so powerful about assuming a null distribution, we can directly quantify the probability of values falling in it.

Being able to quantify this uncertainty allows us to draw conclusions about our data. Were we wrong about our assumptions about that dog breed or is it just simple sample variance? Controlling the size or level of the test allows us to control just how much evidence we need to change our mind as well, giving flexibility to tests that have different levels of risk for a false rejection.

2

How is the significance level related to Type I error? Why do we use hypothesis testing methods that start by assuming the null hypothesis is true instead of initially assuming the alternative is true?

Answer:

The Type I error is directly controlled by the chosen significance level. To risk oversimplifying, they are basically the same thing. You can think of the significance level being a sliding point on a distribution where we decide how far out we need to go to decide to reject the null hypothesis. If we stay closer in, which has a large significance level, there's a higher chance we'll falsly reject the null. We don't need as much evidence, so there's more risk of a false rejection. Going further out requires data to be more extreme and more convincing for us to reject, giving us a lower Type I error.

As for the second part, I kind of answered it in Problem 1. However, a null distribution gives us something we can know for certain. $\theta = 5$ actually gives us a distribution we can examine and compare our data too. If we tried to assume $\theta \neq 5$ where do we start? There are infinitely many distributions that meet that criteria. We wouldn't be able to really do anything with that.

3

What exactly do the numerator and denominator of the likelihood ratio test represent?

Answer:

The numerator represents the likelihood of the null hypothesis and the denominator represents the alternative hypothesis. We put em on top of eachother and directly compare their likelihoods.

4

Based on your answer in question 3, explain what small (close to 0) and large (close to 1) values of the likelihood ratio test statistic mean in terms of the probability of parameter values stated in the null and alternative hypotheses.

Answer:

Small values of the LRT indicate that the likelihood of the data given the null hypothesis is extremely small in comparison to the likelihood of our data given the alternative. That is, the alternative is far more likely for the data than the null. This would indicate we have a lot of evidence to warrant rejecting the null hypothesis.

Values close to 1 indicate that the likelihoods of the null and alternative are nearly identical, thus implying that the data is just as likely in either case. This would mean low evidence towards rejecting the null.

Part 2

8.5

A random sample $X_1, X_2, ..., X_n$is drawn from a pareto population with pdf

$$f(x \mid \theta, v) = \frac{\theta v^\theta}{x^{\theta+1}} I_{[v,\infty)}(x), \quad \theta > 0, \quad v > 0$$

A

Find the MLEs of θ and v.

First we'll start with the likelihood function.

$$L(\theta, v \mid \vec{x}) = \left(\theta v^\theta\right)^n \left(\prod_{i=1}^n x_i^{-(\theta+1)}\right) \left(\prod_{i=1}^n I_{[v,\infty)}(x_i)\right)$$

Maximizing this function with respect to v requires making v as large as possible. However, v has an upper bound of $x_{(1)}$. Therefore, $\hat{v}_{\text{MLE}} = x_{(1)}$.

As for θ, we proceed with the usual workflow.

$$\begin{aligned}
LL(\theta, v \mid \vec{x}) &= n(\ln\theta + \theta \ln v) + \left(\ln \prod x_i^{-(\theta+1)}\right), \quad v \leq x_{(1)} \\
&= n\ln\theta + n\theta\ln v - (\theta+1)\sum \ln(x_i) \\
\frac{d}{d\theta} LL(\theta, v \mid \vec{x}) &= \frac{n}{\theta} + n\ln v - \sum \ln x_i \\
0 &= \frac{n}{\theta} + n\ln v - \sum \ln x_i \\
\frac{n}{\theta} &= \sum \ln x_i - n\ln v \\
\theta &= \frac{n}{\sum \ln x_i - n\ln v} \\
\theta &= \frac{n}{\sum \ln x_i - n\ln x_{(1)}} \\
\frac{d^2}{d\theta^2} LL(\theta, v \mid \vec{x}) &= -\frac{n}{\theta^2} < 0 \; \forall \; \theta
\end{aligned}$$

Therefore,

$$\hat{\theta}_{\text{MLE}} = \frac{n}{\sum \ln x_i - n\ln x_{(1)}}$$

and

$$\hat{v}_{\text{MLE}} = x_{(1)}$$

B

Show that the LRT of $H_0 : \theta = 1$, v unknown. Versus. $H_1 : \theta \neq 1$, v unknown, has critical region of the form

$$\{\vec{x} : T(\vec{x}) \leq c_1 \text{ or } T(\vec{x}) \geq c_2\}$$

where $0 < c_1 < c_2$ and

$$T = \ln\left[\frac{\prod_{i=1}^{n} x_i}{x_{(1)}^n}\right]$$

Before we start, let us examine T and how it relates to $\hat{\theta}_{\text{MLE}}$.

$$\begin{aligned}
\hat{\theta}_{\text{MLE}} &= \frac{n}{\sum \ln x_i - n \ln x_{(1)}} \\
&= \frac{n}{\ln \prod x_i - n \ln x_{(1)}} \\
&= \frac{n}{\ln \prod x_i - \ln x_{(1)}^n} \\
&= \frac{n}{\ln\left(\frac{\prod x_i}{x_{(1)}^n}\right)} \\
&= \frac{n}{T}
\end{aligned}$$

Now we set up the numerator and denominator of the LRT.

$$\begin{aligned}
\lambda(\vec{x}) &= \frac{L(\hat{\theta}_0, \hat{v} \mid \vec{x})}{L(\hat{\theta}, \hat{v} \mid \vec{x})} \\
L(\hat{\theta}_0 = 1, \hat{v} \mid \vec{x}) &= (1v^1)^n \prod_{i=1}^{n} x_i^{-(1+1)}, \quad v \leq x_{(1)} \\
&= x_{(1)}^n \prod_{i=1}^{n} x^{-2} \\
L(\hat{\theta}, \hat{v} \mid \vec{x}) &= \left(\frac{n}{T} x_{(1)}^{n/T}\right)^n \prod_{i=1}^{n} x_i^{-\left(\frac{n}{T}+1\right)} \\
&= \left(\frac{n}{T}\right)^n x_{(1)}^{n^2/T} \prod_{i=1}^{n} x_i^{-\left(\frac{n}{T}+1\right)}
\end{aligned}$$

Now we get into the bulk of the work. Before we dive into the algebra we need to know our goal. We want to get $\lambda(\vec{x})$ into as simple a function of T as possible. All of the rearranging were doing is with that in mind.

$$\begin{aligned}
\lambda(\vec{x}) &= \frac{L(\hat{\theta}_0, \hat{v} \mid \vec{x})}{L(\hat{\theta}, \hat{v} \mid \vec{x})} \\
&= \frac{x_{(1)}^n \prod_{i=1}^n x^{-2}}{\left(\frac{n}{T}\right)^n x_{(1)}^{n^2/T} \prod_{i=1}^n x_i^{-\left(\frac{n}{T}+1\right)}} \\
&= \left(\frac{T}{n}\right)^n x_{(1)}^{n-\frac{n^2}{T}} \left(\prod_{i=1}^n x_i\right)^{\frac{n}{T}-1} \\
&= \left(\frac{T}{n}\right)^n x_{(1)}^{-n(\frac{n}{T}-1)} \left(\prod_{i=1}^n x_i\right)^{\frac{n}{T}-1} \\
&= \left(\frac{T}{n}\right)^n \left(\frac{\prod x_i}{x_{(1)}^n}\right)^{\frac{n}{T}-1} \\
&= \left(\frac{T}{n}\right)^n \exp\left(\left(\frac{n}{T}-1\right) \ln\left(\frac{\prod x_i}{x_{(1)}^n}\right)\right) \\
&= \left(\frac{T}{n}\right)^n \exp\left(\left(\frac{n}{T}-1\right) T\right) \\
&= \left(\frac{T}{n}\right)^n \exp\left(n-T\right) \\
&= \left(\frac{T}{n}\right)^n e^{n-T}
\end{aligned}$$

Now for the critical region. We want to show that the critical region has the form:

$$\{\vec{x} : T(\vec{x}) \leq c_1 \text{ or } T(\vec{x}) \geq c_2\}$$

What we need to acknowledge is that we reject the null hypothesis when $\lambda(\vec{x})$ is small. So we want to find the max of this function so we can figure out where these small values are.

$$\begin{aligned}
\lambda(\vec{x}) &= \left(\frac{T}{n}\right)^n e^{n-T} \\
L\lambda(\vec{x}) &= n \ln T - n \ln n + n - T \\
\frac{d}{dT} &= \frac{n}{T} - 1 \\
0 &= \frac{n}{T} - 1 \\
n &= T \\
\frac{d^2}{dT^2}\lambda(\vec{x}) &= -\frac{n}{T^2} < 0 \,\forall\, T
\end{aligned}$$

So we have $n = T$ is the max of this function. What we can glean from this is our really small values are at the tails of this function. So when T is way smaller or way larger than n. So we have our two sided rejection region. As for how much larger than n it needs to be, we would scale that depending on the size or level test we want. Which gives us our c_1 and c_2. So our critical region has the form:

$$\{\vec{x} : T(\vec{x}) \leq c_1 \text{ or } T(\vec{x}) \geq c_2\}$$

Additional Question

Explain how you could use the fact that $2T$ has a chi-squared distribution with $2n - 2$ degrees of freedom to specify the rejection region for the likelihod ratio test described in part b.

What's great about this is it gives us a straightforward and concrete way to actually get our c_1 and c_2. We can set it up similar to how we would pick our constants with a standard normal or whatever.

So if $2T \sim \chi^2_{2n-2}$, $T \sim \frac{1}{2}\chi^2_{2n-2}$

we can modify our critical region to look like:

$$\left\{\vec{x} : T(\vec{x}) \leq \chi^2_{2n-2,\alpha/2} \text{ or } T(\vec{x}) \geq \chi^2_{2n-2,1-\alpha/2}\right\}$$

So now we have an intuitive and reliable way to get c_1 and c_2 based on some desired α.

8.6

Suppose that we have two independent random samples: $X_1, X_2, ..., X_n$are exponential(θ), and $Y_1, \cdots, Y_m$ are exponential(μ).

A

Find the LRT of $H_0 : \theta = \mu$ versus $H_1 : \theta \neq \mu$.

The key thing to take advantage of here is the fact that X is independent of Y. This makes the joint pdf and joint likelihood way simpler to deal with as it's just the product of their respective functions.

First things first though is figuring out the MLEs of X and Y. What's nice is they're basically identical so we just need to find one of em.

MLES

$$\begin{aligned}
f_X(x \mid \theta) &= \frac{1}{\theta} e^{-x/\theta} \\
L(\theta \mid \vec{x}) &= \theta^{-n} \exp\left(-\theta^{-1} \sum x_i\right) \\
LL(\theta \mid \vec{x}) &= -n \ln(\theta) - \theta^{-1} \sum x_i \\
\frac{d}{d\theta} LL(\theta \mid \vec{x}) &= -\frac{n}{\theta} + \frac{1}{\theta^2} \sum x_i \\
0 &= -\frac{n}{\theta} + \frac{1}{\theta^2} \sum x_i \\
\frac{\sum x_i}{\theta^2} &= \frac{n}{\theta} \\
\frac{1}{\theta^2} &= \frac{n}{\sum x_i} \frac{1}{\theta} \\
1 &= \frac{n\theta}{\sum x_i} \\
\theta &= \frac{1}{n} \sum x_i \\
\hat{\theta} &= \bar{x}
\end{aligned}$$

So our $\hat{\theta} = \bar{x}$. Let's verify that it's a maximum.

$$
\begin{aligned}
\frac{d^2}{d\theta^2} &= \frac{n}{\theta^2} - \frac{2\sum x_i}{\theta^3} \\
&= \frac{n}{\bar{x}^2} - \frac{2n\bar{x}}{\bar{x}^3} \\
&= \frac{n}{\bar{x}^2} - \frac{2n}{\bar{x}^2} \\
&= \frac{n-2n}{\bar{x}^2} \\
&= -\frac{n}{\bar{x}^2} < 0
\end{aligned}
$$

Therefore $\hat{\theta}_{\text{MLE}} = \bar{x}$ and, by extension, $\hat{\mu}_{\text{MLE}} = \bar{y}$.

Let's take a look at our test here.

$$
\begin{aligned}
\lambda(\vec{x}, \vec{y}) &= \frac{L(\theta)L(\mu)}{L(\theta = \hat{\theta})L(\mu = \hat{\mu})} \\
&= \frac{\theta^{-n} e^{-\frac{1}{\theta}\sum x_i} \theta^{-m} e^{-\frac{1}{\theta}\sum y_j}}{\hat{\theta}^{-n} e^{-\frac{1}{\hat{\theta}}\sum x_i} \hat{\mu}^{-m} e^{-\frac{1}{\hat{\mu}}\sum y_j}}
\end{aligned}
$$

So we took care of the denominator, when $\theta \neq \mu$, but what about our numerator? Let's take a look. In this case $\theta = \mu = \phi$.

$$
\begin{aligned}
L(\theta = \phi \mid \vec{x}) \cdot L(\mu = \phi \mid \vec{y}) &= \phi^{-n} \exp\left(-\frac{1}{\phi}\sum x_i\right) \phi^{-m} \exp\left(-\frac{1}{\phi}\sum y_j\right) \\
LL(\theta = \phi \mid \vec{x}) \cdot L(\mu = \phi \mid \vec{y}) &= -n\ln(\phi) - \frac{1}{\phi}\sum x_i m \ln(\phi) - \frac{1}{\phi}\sum y_j \\
&= -\ln(\phi)(n+m) - \frac{1}{\phi}\left(\sum x_i + \sum y_j\right) \\
\frac{d}{d\phi} LL(\theta = \phi \mid \vec{x}) \cdot L(\mu = \phi \mid \vec{y}) &= -\frac{n+m}{\phi} + \frac{\sum x_i + \sum y_j}{\phi^2} \\
\implies \hat{\phi} &= \frac{\sum x_i + \sum y_j}{n+m}
\end{aligned}
$$

Following the same logic as our previous MLE derivation.

LRT

So now that we have our components we can tackle the LRT. Some definitions first.

$$
\begin{aligned}
\sum x_i &= S_X \qquad\qquad\qquad n + m = N \\
\sum y_j &= S_Y \\
S_X + S_Y &= S
\end{aligned}
$$

This is just to save my poor sanity as we proceed through the algebra. Alright. Let's do this. Note that $S_X = \sum x_i = n\bar{x}$.

$$
\begin{aligned}
\lambda(\vec{x}, \vec{y}) &= \frac{L(\theta = \hat{\phi})L(\mu = \hat{\phi})}{L(\theta = \hat{\theta})L(\mu = \hat{\mu})} \\
&= \frac{\hat{\phi}^{-n} e^{-\frac{1}{\hat{\phi}} S_X} \hat{\phi}^{-m} e^{-\frac{1}{\hat{\phi}} S_Y}}{\hat{\theta}^{-n} e^{-\frac{1}{\hat{\theta}} S_X} \hat{\mu}^{-m} e^{-\frac{1}{\hat{\mu}} S_Y}} \\
&= \frac{\left(\frac{S}{N}\right)^{-n} \exp\left(-\frac{N}{S} S_X\right) \left(\frac{S}{N}\right)^{-m} \exp\left(-\frac{N}{S} S_Y\right)}{\bar{x}^{-n} \exp\left(-\frac{1}{\bar{x}} S_X\right) \bar{y}^{-m} \exp\left(-\frac{1}{\bar{y}} S_Y\right)} \\
&= \frac{\left(\frac{S}{N}\right)^{-(n+m)} \exp\left(-\frac{N}{S}\left(S_X + S_Y\right)\right)}{\left(\frac{S_X}{n}\right)^{-n} \exp\left(-\frac{n}{S_X} S_X\right) \left(\frac{S_Y}{m}\right)^{-m} \exp\left(-\frac{m}{S_Y} S_Y\right)} \\
&= \frac{\left(\frac{S}{N}\right)^{-N} \exp\left(-\frac{N}{S} S\right)}{\left(\frac{S_X}{n}\right)^{-n} \left(\frac{S_Y}{m}\right)^{-m} \exp\left(-(n+m)\right)} \\
&= \frac{\left(\frac{S}{N}\right)^{-N} \exp\left(-N\right)}{\left(\frac{S_X}{n}\right)^{-n} \left(\frac{S_Y}{m}\right)^{-m} \exp\left(-N\right)} \\
&= \frac{\left(\frac{S}{N}\right)^{-N}}{\left(\frac{S_X}{n}\right)^{-n} \left(\frac{S_Y}{m}\right)^{-m}}
\end{aligned}
$$

And so we reject H_0 if $\lambda(\vec{x}, \vec{y}) < c$ for some desired α.

B

Show that the test in the previous part can be based on the statistic:

$$T = \frac{\sum X_i}{\sum X_i + \sum Y_i}$$

Okay so this is just more algebra. Using our earlier definitions we're looking for:

$$T = \frac{S_X}{S}$$

So let's get to work.

$$\begin{aligned}
\lambda(\vec{x}, \vec{y}) &= \frac{\left(\frac{S}{N}\right)^{-N}}{\left(\frac{S_X}{n}\right)^{-n}\left(\frac{S_Y}{m}\right)^{-m}} \\
&= \frac{N^N S^{-N}}{S_X^{-n} n^n S_Y^{-m} m^m} \\
&= \frac{N^N}{n^n m^m} \frac{S_X^n S_Y^m}{S^N} \\
&= \frac{N^N}{n^n m^m} \left(\frac{S_X}{S}\right)^n \left(\frac{S_Y}{S}\right)^m \\
&= \frac{N^N}{n^n m^m} T^n \left(\frac{S_Y}{S}\right)^m
\end{aligned}$$

Because $S = S_X + S_Y$,

$$\frac{S_Y}{S} = 1 - \frac{S_X}{S} = 1 - T$$

So,

$$\lambda(\vec{x}, \vec{y}) = \frac{N^N}{n^n m^m} T^n (1 - T)^m$$

And with that we're done.

C

Find the distribution of T when H_0 is true.

Yet again we look at T.

$$T = \frac{S_X}{S_X + S_Y}$$

Some things to consider, what is the distribution of S_X? $S_X + S_Y$? We'll need this information before we can proceed.

Lets start with S_X. Using the mgf of X,

$$M_{\sum x}(t) = \left(\frac{1}{1-\theta t}\right)^n$$

Which is the mgf of a gamma distribution with $\alpha = n, \beta = \theta$. Similar logic for S_Y, but $\alpha = m$ instead. They both share β as we assume H_0 is true for this problem. So, for the distribution of $S_X + S_Y$,

$$M_{S_X+S_Y}(t) = \left(\frac{1}{1-\theta t}\right)^n \cdot \left(\frac{1}{1-\theta t}\right)^m$$

Giving us a gamma distribution with $\alpha = n + m, \beta = \theta$.

With that we can set up a bivariate transformation to get the distribution of T.

$$\begin{aligned} U &= \frac{S_X}{S_X + S_Y} & S_X &= UV & S_Y &= V - S_X \\ V &= S_X + S_Y & & & &= V - VU \\ & & & & &= V(1-U) \end{aligned}$$

$$\begin{aligned} J &= \begin{vmatrix} \frac{dS_X}{dU} & \frac{dS_X}{dV} \\ \frac{dS_Y}{dU} & \frac{dS_Y}{dV} \end{vmatrix} \\ &= \begin{vmatrix} v & u \\ -v & 1-u \end{vmatrix} \\ &= v(1-u) + uv \\ &= v \end{aligned}$$

Now for the actual bulk of the work. We'll get the joint pdf for the transformation and then get the marginal distribution of T.

$$\begin{aligned}
f_{u,v}(u,v) &= F_{S_X,S_Y}(S_X = uv, S_Y = v(1-u)) \cdot |J| \\
&= F_{S_X}(uv)F_{S_Y}(v(1-u))v \\
&= (\Gamma(n)\theta^n)^{-1}(uv)^{n-1}\exp\left(-\frac{uv}{\theta}\right)(\Gamma(m)\theta^m)(v(1-u))^{m-1}\exp\left(-\frac{v(1-u)}{\theta}\right)v \\
&= (\Gamma(n)\Gamma(m)\theta^{n+m})^{-1}(uv)^{n-1}(v(1-u))^{m-1}\exp\left(-\frac{1}{\theta}(uv+v-vu)\right)v \\
&= (\Gamma(n)\Gamma(m)\theta^{n+m})^{-1}u^{n-1}v^{n-1}v^{m-1}(1-u)^{m-1}ve^{-v/\theta} \\
&= (\Gamma(n)\Gamma(m)\theta^{n+m})^{-1}u^{n-1}v^{n+m-1}(1-u)^{m-1}e^{-v/\theta}
\end{aligned}$$

I swear it's not as bad as it looks.

Okay, so now we just get the marginal distribution and pray to the higher beings that watch over the end of the semester that this becomes a distribution we recognize.

$$\begin{aligned}
f_u(u) &= \int_0^\infty (\Gamma(n)\Gamma(m)\theta^{n+m})^{-1}u^{n-1}v^{n+m-1}(1-u)^{m-1}e^{-v/\theta}dv \\
&= (\Gamma(n)\Gamma(m)\theta^{n+m})^{-1}u^{n-1}(1-u)^{m-1}\int_0^\infty v^{n+m-1}e^{-v/\theta}dv
\end{aligned}$$

The integrand there is of the form of an unnormalized gamma distribution with $\alpha = n+m, \beta = \theta$. Thus, the integrand evaluates to the inverse normalizing constant of the gamma distribution, $\Gamma(\alpha)\beta^\alpha$.

$$\begin{aligned}
f_u(u) &= (\Gamma(n)\Gamma(m)\theta^{n+m})^{-1}u^{n-1}(1-u)^{m-1} \cdot \Gamma(n+m)\theta^{n+m} \\
&= \frac{\Gamma(n+m)}{\Gamma(n)\Gamma(m)}u^{n-1}(1-u)^{m-1}
\end{aligned}$$

Which is the pdf of a beta distribution with $\alpha = n, \beta = m$. Therefore,

$$T \sim Beta(n,m)$$

8.18

Let $X_1, X_2, ..., X_n$be as random sample from an $N(\theta, \sigma^2)$ population, with σ^2 known. An LRT of $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$ is a test that rejects H_0 if

$$\frac{|\bar{X} - \theta_0|}{\sigma/\sqrt{n}} > c$$

A

Find an expression, in terms of standard normal probabilities, for the power function of this test.

So our goal here is to look at the power function and see if we can't write it in the form of Z, a standard normal random variable. This problem is primarily just algebraic rearranging.

What this means is that we want $\bar{X} - \theta$ in the numerator there, not θ_0.

$$\begin{aligned}
\beta(\theta) &= P\left(\frac{|\bar{X} - \theta_0|}{\sigma/\sqrt{n}} > c\right) \\
\beta(\theta) &= 1 - P\left(\frac{|\bar{X} - \theta_0|}{\sigma/\sqrt{n}} \leq c\right) && \text{(Flip probability)} \\
&= 1 - P\left(|\bar{X} - \theta_0| \leq \frac{c\sigma}{\sqrt{n}}\right) \\
&= 1 - P\left(-\frac{c\sigma}{\sqrt{n}} \leq \bar{X} - \theta_0 \leq \frac{c\sigma}{\sqrt{n}}\right) && \text{(Undo Absolute Val)} \\
&= 1 - P\left(-\frac{c\sigma}{\sqrt{n}} + \theta_0 \leq \bar{X} \leq \frac{c\sigma}{\sqrt{n}} + \theta_0\right) \\
&= 1 - P\left(-\frac{c\sigma}{\sqrt{n}} + \theta_0 \leq \bar{X} + \theta - \theta \leq \frac{c\sigma}{\sqrt{n}} + \theta_0\right) && \text{(Add } \theta - \theta\text{)} \\
&= 1 - P\left(-\frac{c\sigma}{\sqrt{n}} + \theta_0 - \theta \leq \bar{X} - \theta \leq \frac{c\sigma}{\sqrt{n}} + \theta_0 - \theta\right) \\
&= 1 - P\left(-\frac{c\sigma}{\sqrt{n}} + \theta_0 - \theta \leq (\bar{X} - \theta)\frac{\sigma/\sqrt{n}}{\sigma/\sqrt{n}} \leq \frac{c\sigma}{\sqrt{n}} + \theta_0 - \theta\right) && \text{(Mult by 1)} \\
&= 1 - P\left(-\frac{c(\sigma/\sqrt{n}) + \theta_0 - \theta}{\sigma/\sqrt{n}} \leq \frac{\bar{X} - \theta}{\sigma/\sqrt{n}} \leq -\frac{c(\sigma/\sqrt{n}) + \theta_0 - \theta}{\sigma/\sqrt{n}}\right) \\
&= 1 - P\left(-c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}} \leq Z \leq c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right) \\
&= 1 + P\left(Z \leq -c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right) - P\left(Z \leq c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right) \\
&= 1 + F_Z\left(-c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right) - F_Z\left(c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right)
\end{aligned}$$

Phew. Alright, that's done.

B

The experimenter desires a Type I error probability of .05, and a maximum Type II error probability of .25 at $\theta = \theta_0 + \sigma$. Find values of n, c that will achieve this.

To find c we will examine the power function given the null hypothesis where $\theta = \theta_0$. This simplifies things down quite a bit.

$$\begin{aligned}
\beta(\theta) &= 1 + F_Z(-c) - F_Z(c) \\
0.05 &= 1 + F_Z(-c) - F_Z(c) \\
0.05 &= 1 + (1 - F_Z(c)) - F_Z(c) \\
0.05 &= 2 - 2F_Z(c) \\
-1.95 &= -2F_Z(c) \\
0.975 &= F_Z(c) \\
c &= F_Z^{-1}(0.975) \\
c &= 1.96
\end{aligned}$$

Important note here that we used the inverse norm. We could've done this using calculus but it is the final homework so I used my calculator for this. This gives us the constant we need given a specific probability.

Now, for n we need to look at the Type II error rate. Here we'll substitute $\theta = \theta_0 + \sigma$.

Really important to point out here that a Type II error rate of 0.25 gives us $\beta(\theta \mid H_1) = 1 - P(\text{Type II error}) = 1 - 0.25 = 0.75$.

$$\begin{aligned}
\beta(\theta_0 + \sigma) &= 1 + F_Z\left(-c + \frac{\theta_0 - \theta_0 - \sigma}{\sigma/\sqrt{n}}\right) - F_Z\left(c + \frac{\theta_0 - \theta_0 - \sigma}{\sigma/\sqrt{n}}\right) \\
&= 1 + F_Z\left(-c + \frac{-\sigma}{\sigma/\sqrt{n}}\right) - F_Z\left(c + \frac{-\sigma}{\sigma/\sqrt{n}}\right) \\
&= 1 + F_Z\left(-c - \sqrt{n}\right) - F_Z\left(c - \sqrt{n}\right) \\
0.75 &= 1 + F_Z\left(-1.96 - \sqrt{n}\right) - F_Z\left(1.96 - \sqrt{n}\right)
\end{aligned}$$

Note here that $F_Z(-1.96 - \sqrt{n})$ is going to be approximately 0 for nearly any n. For example, $n = 10$ already puts this value at $1.5e^{-7}$. So we can simplify this further.

$$\begin{aligned}
0.75 &= 1 - F_Z\left(1.96 - \sqrt{n}\right) \\
-.25 &= -F_Z\left(1.96 - \sqrt{n}\right) \\
.25 &= F_Z\left(1.96 - \sqrt{n}\right) \\
F_Z^{-1}(.25) &= 1.96 - \sqrt{n} \\
-.675 &= 1 - \sqrt{n} \\
n &= 6.94 \\
n &\approx 7
\end{aligned}$$

So we would need $c = 1.96$ and $n = 7$ to achieve these results.