1 Large Sample Testing

- As Statisticians, the most common testing situation we will face are situations in which the Central Limit theorem applies
- The typical structure will be along the following lines:

$$\begin{aligned} H_0 : \theta &= \theta_0 \\ H_a : \theta &= \theta_a (> \theta_0) \\ U &= \bar{X} \\ RR &= \{\bar{X} > k\} \end{aligned}$$

- We can think of the RR as $= \{\bar{x} : \bar{x} > k\} = \{\bar{x} : \frac{\bar{x}-\theta_0}{\widehat{\sigma_X}/\sqrt{n}} > \frac{k-\theta_0}{\widehat{\sigma_X}/\sqrt{n}}\}$, where σ_X^2 is the variance of the underlying distribution
- We also know that under the CLT, if our sample is large enough $\frac{\bar{x}-\theta_0}{\widehat{\sigma_X}/\sqrt{n}}$ is approximately distributed $N(0,1)$. This is assuming that $E[X_1] = \theta_0$
- This means that

$$\begin{aligned} P(\text{Type I Error}) &= P(\bar{X} \in RR|\theta = \theta_0) \\ &= P(\bar{X} \in \{\bar{x} : \frac{\bar{x}-\theta_0}{\widehat{\sigma_X}/\sqrt{n}} > \frac{k-\theta_0}{\widehat{\sigma_X}/\sqrt{n}}\}|\theta = \theta_0) \\ &= P(\frac{\bar{X}-\theta_0}{\widehat{\sigma_X}/\sqrt{n}} > \frac{k-\theta_0}{\widehat{\sigma_X}/\sqrt{n}}\}|\theta = \theta_0) \\ &\approx P(Z > \frac{k-\theta_0}{\widehat{\sigma_X}/\sqrt{n}}) \leftarrow \text{Where Z} \sim N(0,1) \end{aligned}$$

- So, if we select $k = \theta_0 + Z_\alpha \widehat{\sigma_X}/\sqrt{n}$ then our significance will be approximately $P(Z > \frac{(\theta_0 + Z_\alpha \widehat{\sigma_X}/\sqrt{n})-\theta_0}{\widehat{\sigma_X}/\sqrt{n}}) = P(Z > Z_\alpha) = \alpha$
- This means we can select our rejection region for a fixed significance level when we have a large sample without known the underlying distribution (approximately).

1.1 Example

Let $X_1, ..., X_n$ be i.i.d $Bernoulli(p)$. Suppose that we want to test $H_0 : p = p_0$ vs. $H_a : p > p_0$. Then we know that under the null hypothesis $\widehat{\sigma_X}/\sqrt{n} = \sqrt{p_0(1-p_0)/n}$ and that $\frac{\bar{X}-p_0}{\sqrt{p_0(1-p_0)/n}}$ is approximately distributed $N(0,1)$. This means if we want to conduct a hypothesis test with a significance of $\alpha = .05$ we can have $U = \frac{\bar{X}-p_0}{\sqrt{p_0(1-p_0)/n}}$ and our Rejection Region would be $RR = \{Z > Z_{.05}\}$. Because of the CLT, we can confirm that our significance is indeed .05.

$$\begin{aligned}
P(\text{type I error}) &= P(U \in RR|H_0 \text{ is true}) \\
&= P(U > Z_{.05}|p = p_0) \\
&\approx P(Z > Z_{.05}) \leftarrow \text{ where } Z \sim N(0,1) \\
&= .05
\end{aligned}$$

1.2 Calculating β and finding the sample size

- To calculate the probability of a type II error (β) in large sample testing situations is fairly stragihtforward.
- to determine this, we return to our original situation:

$$\begin{aligned}
H_0 : \theta &= \theta_0 \\
H_a : \theta &= \theta_a \\
U &= \bar{X} \\
RR &= \{\bar{X} > k\}
\end{aligned}$$

- We now know that we would begin by fixing α. From this point we would then find k. In the large sample cases this would be $k = \theta_0 + Z_\alpha \widehat{\sigma_X}/\sqrt{n}$
- This means that if we want to know β for the case when $\theta = \theta_a > \theta_0$ we see that

$$\begin{aligned}
P(\text{Type II Error}) &= P(\bar{X} \notin RR|\theta = \theta_a > \theta_0) \\
&= P(\bar{X} \leq k|\theta = \theta_a) \\
&= P(\frac{\bar{X}-\theta_a}{\widehat{\sigma_X}/\sqrt{n}} \leq \frac{k-\theta_a}{\widehat{\sigma_X}/\sqrt{n}}|\theta = \theta_a) \\
&\approx P(Z \leq \frac{k-\theta_a}{\widehat{\sigma_X}/\sqrt{n}}) \leftarrow \text{ Where } Z \sim N(0,1)
\end{aligned}$$

- With this formula for β we can determine how large of a smaple size we need in order to fix both α and β for a given θ_a
- We already have

$$\begin{aligned}
P(\text{type I error}) &= P(\bar{X} \in RR|H_0 \text{ is true}) \\
&= P(\frac{\bar{X}-\theta_0}{\widehat{\sigma_X}/\sqrt{n}} > \frac{k-\theta_a}{\widehat{\sigma_X}/\sqrt{n}}|p = p_0) \\
&\approx P(Z > Z_\alpha) \leftarrow \text{ where } Z \sim N(0,1)
\end{aligned}$$

$$\begin{aligned}
\Rightarrow P(\text{Type II Error}) &= P(\bar{X} \notin RR|\theta = \theta_a > \theta_0) \\
&= P(\bar{X} \leq k|\theta = \theta_a) \\
&= P(\frac{\bar{X} - \theta_a}{\widehat{\sigma_X}/\sqrt{n}} \leq \frac{k - \theta_a}{\widehat{\sigma_X}/\sqrt{n}}|\theta = \theta_a) \\
&\approx P(Z \leq \frac{k - \theta_a}{\widehat{\sigma_X}/\sqrt{n}}) \leftarrow \text{ Where } Z \sim N(0,1) \\
&= P(Z \leq -Z_\beta)
\end{aligned}$$

- These two formulas give us two equations

$$\frac{k - \theta_0}{\widehat{\sigma_X}/\sqrt{n}} = Z_\alpha$$

$$\frac{k - \theta_a}{\widehat{\sigma_X}/\sqrt{n}} = -Z_\beta$$

- Solving for k, and setting the equations equal we have

$$k = \theta_0 + Z_\alpha(\widehat{\sigma_X}/\sqrt{n}) = \theta_a - Z_\beta(\widehat{\sigma_X}/\sqrt{n})$$

which, if we then solve for n we get

$$n = \frac{(Z_\alpha + Z_\beta)^2(\widehat{\sigma_X})^2}{(\theta_a - \theta_0)^2}$$

1.3 Examples

1. Suppose $X_1, ..., X_n$ are i.i.d $N(\mu_x, \sigma_x^2)$ and $Y_1, ..., Y_m$ are i.i.d $N(\mu_y, \sigma^2)$. Suppose we want to test to see if the mean of the first population is larger than the mean of the second population at the α level. In other words we are testing $H_0 : \mu_x = \mu_y$ vs. $H_a : \mu_x > mu_y$. Assuming that the popluations are both sufficiently large (and independent from each other), construct an approximate hypothesis test at the α significance level.
Solution:

We can reformulate our hypotheses as $H_0 : D = \mu_x - \mu_y = 0$ vs $H_a : D = \mu_x - \mu_y > 0$, and we know that with $\hat{\sigma_x^2} = S_x^2$ and $\hat{\sigma_y^2} = S_y^2$ We have that

$$U = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}}$$

is approximately normally distibuted. This means that our rejection region will be $RR = \{u > Z_\alpha\}$ and we can confirm that our significance will be α:

$$\begin{aligned} P(\text{Type I Error}) &= P(U \in RR | H_0 \text{ is True}) \\ &= P(U > Z_\alpha | D = 0) \\ &\approx P(Z > Z_\alpha) \\ &= \alpha \end{aligned}$$

2. if we assume that $m = n$ then in order to fix $\alpha = \alpha_0$ and $\beta = \alpha_0$ when $D = D_a > 0$, how big would n need to be?
Solution:
Using our formula where $\widehat{\sigma_X^2} = S_x^2 + S_y^2$,$\alpha = \beta = \alpha_0$, $\theta_0 = D_0 = 0$ and $\theta_a = D_a > 0$, we have

$$\begin{aligned} n &= \frac{(Z_\alpha + Z_\beta)^2(\widehat{\sigma_X})^2}{(\theta_a - \theta_0)^2} \\ &= \frac{(2Z_{\alpha_0})^2(S_x^2 + S_y^2)}{(D_a)^2} \end{aligned}$$

2 More Hypothesis Testing

2.1 Two Sided Testing

- Sometimes, when we wantto conduct a hypothesis test, we simply want to know if there is evidence that a parameter of interest is a specific value or not.
- We characterize such a test with the following hypothesis statement:

$$H_0 : \theta = \theta_0 \text{ vs } H_a : \theta \neq \theta_0$$

- In this testing situation, it make sense that we want to reject our null hypothesis if we receive information that suggests the true value of θ is too high or too low.
- This means that we will want to reject if our testing statisic is too big or too small.
- This means that the rejection region will be of the form $RR = \{x : x > k_1 \text{ or } x < k_2\}$
- k_1 and k_2 are selected such that we can fix our probability of a type I error to be α
- Typically, we want to make the probabiltiy of a type I error "In either direction" to to be equal. That is we want

$$P(\text{Type I error}) = P(U \in RR | H_0 \text{ is true})$$

$$
\begin{aligned}
&= P(U > k_1 \text{ or } U < k_2|\theta = \theta_0) \\
&= P(U > k_1|\theta = \theta_0) + P(U < k_2|\theta = \theta_0) \\
&= \alpha_1 + \alpha_2 \\
&= \alpha/2 + \alpha/2 \\
&= \alpha
\end{aligned}
$$

- So, to construct our rejection region we simply set

  $$P(U > k_1|\theta = \theta_0) = P(U < k_2|\theta = \theta_0) = \alpha/2$$

  and solve for k_1 and k_2 seperately.

- Generically, our solutions will be $k_1 = U_{\alpha/2}$ and $k_2 = U_{1-\alpha/2}$, where $U_{1-\alpha/2}$ and $U_{\alpha/2}$ were the $\alpha/2$ and $1 - \alpha/2$ percentiles of the distribution (or approximate distrubution) of U

2.2 CIs and Hypothesis testing

- When we constructed our CIs, we began by constructing a pivot (or an approximate pivot) U
- The, we said we wanted to find a, b such that the interval (a, b) contained the middle $1 - \alpha$ of the distribution of U.
- We wrote this in an equation as $P(U_{1-\alpha/2} < U < U_{\alpha/2}) = 1 - \alpha$, where $U_{1-\alpha/2}$ and $U_{\alpha/2}$ were the $\alpha/2$ and $1-\alpha/2$ percentiles of the distribution (or approximate distrubution) of U
- Note, that if we use U to be our test statisic, then our rejection region will be $RR = \{x : x > U_{\alpha/2} \text{ or } x < U_{1-\alpha/2}\}$ This in fact the compliment of our CI!

So, what does this mean?

- We can think of the testing statistic as a function of just θ_0 once the data has been observed
- When we re-write the CI in the form $(\hat{\theta}_L, \hat{\theta}_U)$, we can think of This intveral as the range of posible values of θ_0 that would generate a statistic U (assuming fixed data) that would not have ben rejected by our test

2.3 p-value

Another way a test can be conducted is through the use of a p-value
Definition:
The *p-value* of a given test is the smallest significance for which the test would reject for an observed test statistic

- In other words, it is the probability of observing a statistic as extreme (or more exteme) as what we observed, given that the null hypothesis is true.
- Consider the testing situation when we have $H_0 : \theta = \theta_0$ vs. $H_a : \theta > \theta_0$
- Let X be a random variable from the distribution of U. Let U_0 be the value of the test statistic after observing data $X_1, ..., X_n$. Then the p-value will be

$$\text{p-value} = P(X \geq U_0)$$

- If the alternative hypothesis was instead $H_a : \theta < \theta_0$, then the p-value would be

$$\text{p-value} = P(X \leq U_0)$$

3 Problems Day

3.1 problems

1. Let $X \sim U(0, \theta)$, and consider the hypothesis testing situation $H_0 : \theta = 5$ vs. $H_a : \theta = 2$ with X as the testing statistic
   a) Find α when the the rejection region is $RR = \{x : x < 1.5\}$
   b) Find β when the rejection region is $RR = \{x : x < 1.5\}$
   c) What rejection region should we use if we want to fix $\alpha = .1$?
   d) Suppse we observe $X = 1$. What would be the p-value of this observed statistic value?
2. Let $X \sim Geo(p)$, and consider the hypothesis testing situation $H_0 : p = .2$ vs. $H_a : \theta = .1$ with X as the testing statistic
   a) Show that the cdf of X is $F(x) = 1 - (1 - p)^x$
   b) Find α when the the rejection region is $RR = \{x : x > 2\}$
   c) Find β when the rejection region is $RR = \{x : x > 2\}$
   d) What rejection region should we use if we want to fix $\beta = .271$?
   e) Suppse we observe $X = 3$. What would be the p-value of this observed statistic value?
3. Let $X_1, ..., X_n$ be i.i.d. $poisson(\lambda)$. Suppose that we want to conduct an approximate hypothesis test at the significance level α for the testing situation $H_0 : \lambda = 1$ vs $H_a : \lambda > 1$
   a) Construct an appropriate testing statistic (remember we want CLT to apply under the null hypothesis)
   b) What would the appropriate rejection region be for this hypothesis?

c) Suppose that you observe a sample mean of 1.165 from a sample of $n = 100$. Conduct the test at the $\alpha = .025$ significance level.

d) What would the approximate p-value of your observed test statistic be?