Outlier

Section Summary | Introduction to Statistics
Hypothesis Testing With One Sample

Hypothesis Testing, Type I and Type II Errors

1. Hypothesis testing
    1. When approaching a statistical question, you assume the "Null Hypothesis", which is the assertion that the observed phenomenon happened by pure chance
    2. If we believe that there are other causes besides pure chance, we propose the "Alternative Hypothesis"
    3. In hypothesis testing we test the likelihood that the null hypothesis is true
    4. The results of hypothesis testing are we can either reject or fail to reject the null hypothesis (we never accept the null hypothesis)
    5. When we reject the null hypothesis, we reject it in favor of the alternative hypothesis

2. Type I () errors
    1. A type I error occurs when we reject the null hypothesis but the null hypothesis turns out to be true
    2. This is the same as the from confidence intervals
    3. gives the probability that our confidence interval does not contain the true parameter
    4. It also gives the probability of making type I errors
    5. =0.05 is a common convention
    6. Type I errors are also known as false positives
    7. You choose based on how many mistakes you're willing to accept in the experiment

3. Type II () errors

1. A type II error occurs when we fail to reject the null hypothesis but the null hypothesis is false
2. These are also known as false negatives
3. represents the probability of getting a false negative
4. The probability of rejecting the null hypothesis when the alternative hypothesis is true is known as “power” and given as 1-

Hypothesis Testing, Type I and Type II Errors

1. Null Hypothesis: H_0
   a. The null hypothesis states an assumption on a population parameter.
   b. If the assumption is that the population mean is equal to some number, then the null hypothesis is stated as $H_0: \mu = \mu_0$
   c. The null hypothesis $H_0: \mu \leq \mu_0$is used when the population mean is assumed to be less than μ_0 which is some number.
   d. When the assumption is that the population mean is greater than or equal to some number, then the null hypothesis is stated as $H_0: \mu \geq \mu_0$ where μ_0 is some number.
   e. When testing a population proportion, p is used instead of μ and p represents the population proportion.
   f. The null hypothesis for testing a population proportion can be stated as $H_0: p = p_0$, $H_0: p \leq p_0$, or $H_0: p \geq p_0$ where p_0 represents some number between 0 and 1.
2. Alternative Hypothesis: H_a
   a. The alternative hypothesis states a competing view of the population parameter using the same value from the null.
   b. The alternative hypothesis is determined by the specific problem being dealt with.
   c. The alternative hypothesis $H_a: \mu \neq \mu_0$ means that the population mean is not equal to the assumed number. This is a two-tailed test
   d. The alternative hypothesis $H_a: \mu > \mu_0$ means that the population mean is larger than the assumed number. This is a right-tailed test

e. The alternative hypothesis $H_a: \mu < \mu_0$ means that the population mean is less than the assumed number. This is a left-tailed test

f. When $H_0: \mu \le \mu_0$ there is only one option for the alternative hypothesis: $H_a: \mu > \mu_0$

g. When $H_0: \mu \ge \mu_0$ there is only one option for the alternative hypothesis: $H_a: \mu < \mu_0$

h. The alternative hypothesis for testing a population proportion p can be stated as $H_a: p \neq p_0$, $H_a: p < p_0$, or $H_a: p > p_0$ where p_0 represents some number between 0 and 1.

3. Test Statistic

a. Summarizes the data that is collected

b. There are 3 different test statistics that are used for 1 Sample Hypothesis Tests.

c. $Z = \dfrac{\underline{X} - \mu_0}{\sqrt{\frac{\sigma^2}{n}}}$ is used when testing a population mean with a known population variance σ^2. The sample size is given by n, $\underline{X}$ represents the sample mean, and μ_0 represents the assumed population mean from the null hypothesis.

d. $t = \dfrac{\underline{X} - \mu_0}{\sqrt{\frac{s^2}{n}}}$ is used when testing a population mean with an unknown population variance σ^2. The sample variance s^2 is used instead of the population variance. The sample size is denoted by n, $\underline{X}$ represents the sample mean, and μ_0 represents the assumed population mean from the null hypothesis.

Outlier

e. $Z = \dfrac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ is used when testing a population proportion. The sample size is denoted by n, $\hat{p}$ represents the sample proportion, and p_0 represents the assumed population proportion from the null hypothesis.

4. Rejection Region
   a. This is the collection of all possible test statistic values for which the null hypothesis will be rejected.
   b. If the test statistic falls into the rejection region, we reject the null hypothesis. If the test statistic does not fall into the rejection region, we fail to reject the null hypothesis.
      - We never accept the null hypothesis; we simply fail to reject the null hypothesis.
   c. The region is determined by the alternative hypothesis and the level of significance α.
   d. The level of significance α determines the critical value(s) that determine the border(s) for the rejection region.
   e. The critical value(s) are also determined by which test statistic is used. Using the Z-test statistic (whether for population mean or proportion) means the critical value(s) will come from the standard normal distribution. Using the t test statistic means that the critical value(s) will come from the t-distribution with n-1 degrees of freedom (n is the sample size).
   f. If $H_a: \mu > \mu_0$ or $H_a: p > p_0$, then a right-tailed test is being used and the critical value is the number with α area to the right and $1 - \alpha$ area to the left. Otherwise known as the $(1 - \alpha)*100\%$ percentile (Example: if $\alpha = .01$, then the critical value is the 99th percentile of the distribution)
   g. If $H_a: \mu < \mu_0$ or $H_a: p < p_0$, then a left-tailed test is being used and the critical value is the number with α area to the left and $1 - \alpha$ area to the

right. Otherwise known as the $\alpha * 100\%$ percentile (Example: if $\alpha = .01$, then the critical value is the 1st percentile of the distribution)

h. If $H_a: \mu \neq \mu_0$ or $H_a: p \neq p_0$, then a two-tailed test is being used and alpha is divided up evenly into two tails. There are two critical values. The larger one will have $1 - \frac{\alpha}{2}$ area to the left. So, if $\alpha = .05$, then the critical value is the 97.5th percentile of the distribution. The other critical value is just the opposite of the larger one (example: 1.96 and -1.96).

i. The p-value is the probability of observing a test statistic value more extreme given that the null hypothesis is true.

j. The p-value is an equivalent method to the rejection region. If the p-value is smaller than α, then the test statistic is in the rejection region and we reject the null hypothesis. If the p-value is larger, then the test statistic is not in the rejection region and we fail to reject the Null Hypothesis.

k. How the p-value is calculated is dependent on the distribution of the test statistic (t-distribution for t test statistic and normal distribution for z test statistics) and the alternative hypothesis.

- If $H_a: \mu > \mu_0$ or $H_a: p > p_0$ then the p-value is the probability of observing a value larger than the test statistic (area to the right).
- If $H_a: \mu < \mu_0$ or $H_a: p < p_0$ then the p-value is the probability of observing a value smaller than the test statistic (area to the left).
- If $H_a: \mu \neq \mu_0$ or $H_a: p \neq p_0$ then the p-value is twice the probability of observing a value smaller than the test statistic if the test statistic is negative. The p-value is twice the probability of observing a value larger than the test statistic if the test statistic is positive.

Applying Hypothesis Testing: Two Examples

1. Newborn Birth Weights

   g. Data contains sex and weight for 44 newborns

h. Sample has a mean of 3276 grams.

i. The null and alternative hypothesis are: H_0: $\mu = 3100$ and H_a:$\mu > 3100$

j. With a sample size of 44 we can use the Central Limit Theorem.

k. The population variance is unknown; therefore, the sample variance will be used in its place which means a t-test statistic $t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$ will be used. $\bar{X}$ is the sample mean (it is equal to 3275.955) from the data and μ_0 is the assumed mean from the null hypothesis (it has a value of 3100). The sample standard deviation is s = 528.032. The sample size n in 44.

l. Putting this together, the test statistic is

$$t = \frac{3275.955 - 3100}{\frac{528.032}{\sqrt{44}}} = 2.2104$$

m. Using $\alpha = 0.05$, the t-distribution with 43 degrees of freedom, and the alternative hypothesis (right-tailed test); the critical number is a number with 5% above and 95% below.

n. The critical value is about 1.68 (the 95th percentile of this distribution). This means the rejection region is all values above 1.68.

o. Since 2.2104 is larger than 1.68, the test statistic is in the rejection region and the null hypothesis is rejected.

p. The p-value for this test, will be the probability of observing t test statistic greater than 2.2104. This value is 0.01622

2. Proportion of Children who are Girls

a. Using the same data set as before, sex of the newborn was tracked. 26 of the babies were boys and 18 were girls. This data will be used to test the null hypothesis that the overall proportion of girls born is 0.5 ($H_0: p = 0.5$) versus the alternative hypothesis ($H_a: p \neq 0.5$).

b. The level of significance will be 5% ($\alpha = 0.05$)

c. Assumptions

- The data are independent and identically distributed.
- $np > 10$ and $n(1 - p) > 0$ (this is checked by plugging in the sample size and the assumed proportion from the null hypothesis)

d. With these assumptions, the normal distribution can be used. The test statistic for testing population proportion is $z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$ where

$\hat{p} = \frac{18}{44}$ is the sample proportion, $p_0 = 0.5$ is the assumed population proportion from the null hypothesis, and $n = 44$ is the sample size.

e. Plugging those numbers in gives

$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}} = \frac{\frac{18}{44} - 0.5}{\sqrt{\frac{0.5(1 - 0.5)}{44}}} = -1.206$ which is the value of the test statistic for this test.

f. Since $H_a: p \neq 0.5$ and $\alpha = 0.05$, there are two tails in this test with alpha split between them. Due to symmetry, alpha is split evenly two critical values are needed so that 95% is between them and 5% is in both tails (2.5% in each tail). The upper critical value then will have 95% + 2.5% below it. Making the upper critical value the 97.5 percentile. This value is 1.96 and the other is -1.96 due to symmetry.

g. The rejection region is all of the values greater than 1.96 and all of the values below -1.96. With $z = -1.206$, neither is the case; so "fail to reject the null hypothesis" is the conclusion.

h. The p-value is going to be 2 times the area to the left of -1.206 as this is a two tailed test and the p-value is the probability of observing a test statistic less than -1.206 or a test statistic greater than 1.206. The p-value is $2*P(Z < -1.206) = 2(0.1139) = 0.2278$