

1. In an experiment to see how hypertension is related to smoking habits, the following data was taken on 111 individuals. Test the hypothesis that the proportions of hypertension in the two groups are not equal. Use a .05 level of significance. State the null hypothesis, the test statistic, why your statistic is appropriate, the rejection region, your computation and conclusion.

	Smoking Habits		Total
	Moderate Smokers	Heavy Smokers	
Hypertension	36	30	66
No Hypertension	26	19	45
Total	62	49	111

Let X_1 be the number of hypertensive individuals among n_1 moderate smokers and X_2 the number of hypertensives among n_2 heavy smokers. Thus $X_i \sim \text{Bin}(n_i, p_i)$, where p_i is the population proportion of hypertensives in the i -th group. The null and alternative hypotheses are

$$\begin{aligned} \mathcal{H}_0 &: p_1 = p_2; \\ \mathcal{H}_a &: p_1 \neq p_2. \end{aligned}$$

Because the proportions are equal in the null hypothesis, the appropriate statistic has pooled estimate for p_i . Since we have large samples: $n_1 p_1 = 36 > n_1 q_1 = 26 > 10$ and $n_2 p_2 = 30 > n_2 q_2 = 19 > 10$, by rule of thumb, the z -test is appropriate.

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \text{where } p_i = \frac{X_i}{n_i} \text{ and } \bar{p} = \frac{X_1 + X_2}{n_1 + n_2}.$$

The null hypothesis is rejected if $|Z| \geq z_{\alpha/2}$. For $\alpha = .05$, $z_{.025} = 1.960$. We find

$$\bar{p} = \frac{66}{111}, \quad Z = \frac{\frac{36}{62} - \frac{30}{49}}{\sqrt{\frac{66}{111} \cdot \frac{45}{111} \left(\frac{1}{62} + \frac{1}{49}\right)}} = -.337.$$

Since $|Z| < z_{\alpha/2}$, we cannot reject the null hypothesis. The evidence is not strong that the proportion of hypertensives among moderate smokers is different than the proportion of hypertensives among heavy smokers.

2. The Wanship Widgets Company is interested in evaluating its current inspection procedure on shipments of 50 identical widgets. The procedure is to take a sample of 5 from the 50 and pass the shipment if no more than two are found to be defective. What is the probability that the shipment will be accepted even though actually 20% of the widgets in the shipment are defective?

Let X be the number of defective widgets in the sample of $n = 5$ taken from $N = 50$ without replacement. Suppose the shipment has M defective widgets. Thus X is a hypergeometric variable with $X \sim \text{Hypergeometric}(n = 5, N = 50, M = pN)$, where p is the proportion of defectives in the shipment.

If $p = .2$ then $M = 10$. The probability of accepting the shipment is

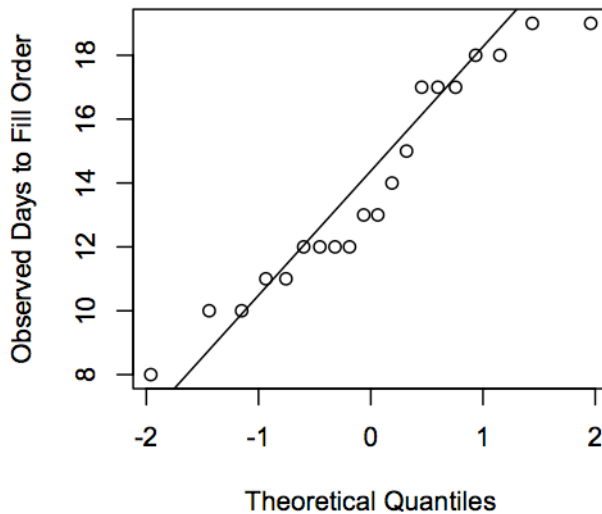
$$\begin{aligned} \beta &= P(X \leq 2) = \text{hyp}(0, 5, 50, 10) + \text{hyp}(1, 5, 50, 10) + \text{hyp}(2, 5, 50, 10) \\ &= \frac{\binom{10}{0} \binom{40}{5}}{\binom{50}{5}} + \frac{\binom{10}{1} \binom{40}{4}}{\binom{50}{5}} + \frac{\binom{10}{2} \binom{40}{3}}{\binom{50}{5}} = .952. \end{aligned}$$

3. Extensive records of Wanship Widgets Co. indicates the number of days to fill orders for their widget controller were

13 18 10 12 19 11 18 17 14 15 17 10 12 8 12 12 17 13 19 11

Does the data strongly indicate that such orders take longer than 12 days to fill? State the null and alternative hypothesis. State the test statistic and rejection region. What assumptions are you making about the data and how well are they satisfied? Perform the computation and state your conclusion. [Hint: $\bar{X} = 13.9$ and $S = 3.354$.]

Normal Q-Q Plot



Let X be the number of days to fill the controller order. Let μ be the population mean of the number of days to fill the order. The null and alternative hypotheses are

$$\begin{aligned} \mathcal{H}_0 &: \mu = 12; \\ \mathcal{H}_a &: \mu > 12. \end{aligned}$$

The test statistic is

$$T = \frac{\bar{X} - 12}{S/\sqrt{n}}.$$

Since we have a small sample ($n = 12 \leq 40$) and the population standard deviation σ is unknown, we use the t -test. The assumption we make in this problem is that the sample comes from a normal distribution. The normal PP-plot shows that the points line up pretty well so there is no strong indication that the distribution is not normal, although there is an “S” shape indicating heavy tails. (In fact the data was generated by a random t -distribution!) The null hypothesis is rejected at the α level of significance if $t \geq t_{\nu, \alpha}$ where $\nu = n - 1 = 19$ is degrees of freedom. We compute using the hint

$$t = \frac{13.9 - 12}{3.354/\sqrt{20}} = 2.533.$$

From table A.8, for $\nu = 19$ d.f., $P(T > 2.5) = .011$ and $P(T > 2.6) = .009$. Thus, interpolating, $P = P(T > 2.533) = .010$. From table A.5, $t_{19, .025} = 2.093$ and $t_{19, .01} = 2.539$ so $.01 < P < .025$. Thus at the $\alpha = .011$ significance level, we reject the null hypothesis: there is strong evidence that it takes longer than 12 days to fill the controller order.

4. Let X denote the number of years that a Wanship Widget operates before it ceases to function. Assume that the probability density function (pdf) is

$$f(x) = \begin{cases} \frac{1}{3}e^{-x/3}, & \text{if } x \geq 0; \\ 0, & \text{if } x < 0. \end{cases}$$

Find the cumulative density function (cdf) for X . Find the probability that the widget operates at least four years. Find the probability that the widget continues to operate at least another four years given that it has been functioning for two years.

Observe that X is an exponential variable $X \sim \text{Exponential}(\lambda = \frac{1}{3})$. The cdf is by definition

$$F(x) = \int_{-\infty}^x f(x) dx = \begin{cases} \int_0^x \frac{1}{3}e^{-x/3} dx = 1 - e^{-x/3}, & \text{if } x \geq 0; \\ 0, & \text{if } x < 0. \end{cases}$$

The probability that the widget lasts at least four years is

$$P(X \geq 4) = 1 - P(X < 4) = 1 - F(4) = 1 - (1 - e^{-4/3}) = .264.$$

The probability that the widget continues to operate at least another four years given that it has been functioning for two years is by the conditional probability formula

$$\begin{aligned} P(X \geq 6 \mid X \geq 2) &= \frac{P(\{X \geq 6\} \cap \{X \geq 2\})}{P(X \geq 2)} = \frac{P(X \geq 6)}{P(X \geq 2)} = \frac{1 - P(X < 6)}{1 - P(X < 2)} \\ &= \frac{1 - F(6)}{1 - F(2)} = \frac{1 - (1 - e^{-6/3})}{1 - (1 - e^{-2/3})} = e^{-4/3} = .264. \end{aligned}$$

Equality in (b.) and (c.) is known as the “memoryless property” of exponential variables.

5. A random sample of Dan’s customers shows that of 30 shoppers, 21 of them regularly used cents-off coupons. Construct a 99% one-sided upper confidence interval for the proportion of customers who use coupons. Finish the following statement that describes your confidence interval “There is a 99% chance that...”

Let X be the number of coupon users. $X \sim \text{Binomial}(n = 21, p)$, where p is the population proportion of coupon users. Since we have a small sample, $n\hat{q} = 9 < 10$, we cannot use the traditional CI given by equation 7.11. Instead we use the unrestricted one-sided upper confidence bound

$$p < \frac{\hat{p} + \frac{z_\alpha^2}{2n} + z_\alpha \sqrt{\frac{\hat{p}\hat{q}}{n} + \frac{z_\alpha^2}{4n^2}}}{1 + \frac{z_\alpha^2}{n}}$$

Since $z_{.01} = 2.326$, $\hat{p} = 21/30 = .7$, $\hat{q} = .3$, $n = 30$ we find

$$p < \frac{.7 + \frac{2.326^2}{60} + 2.326 \sqrt{\frac{(.7)(.3)}{30} + \frac{2.326^2}{3600}}}{1 + \frac{2.326^2}{30}} = .851.$$

The following statement describes the interval.

There is a 99 % chance that our interval has captured p.

The statement “There is a 99% chance that $p < .851$.” is vague and imprecise. On the one hand it can be interpreted as my statement, but it can also be interpreted as “.99 = $P(p < .851)$ ” which is false because although in this statement we intend that “.851” be the random variable, it is a fixed number.

6. Suppose that 10% of all shafts used in widget construction are nonconforming. A group of 200 shafts is randomly selected. What is the exact probability that not more than 30 are nonconforming? Give the formula (as a sum) but do not evaluate. Compute the probability using the best approximation. Why is your approximation valid?

Let X denote the number of nonconforming shafts in the sample of $n = 200$. $X \sim \text{Binomial}(n = 200, p = .10)$, where p is the population proportion of nonconforming shafts. Then the probability that not more than 30 are nonconforming is

$$P = P(X \leq 30) = \text{Binomial}(30, 200, .1) = \sum_{i=0}^{30} \binom{200}{i} (.1)^i (.9)^{200-i}.$$

The rule of thumb for Poisson approximation is not satisfied, although the rule for normal approximation is: $nq > np = 200 \cdot (.1) = 20 > 10$. The best approximation includes the continuity correction

$$P(X \leq x) \approx \Phi\left(\frac{x + .5 - np}{\sqrt{npq}}\right)$$

Thus we find

$$P(X \leq 30) \approx \Phi\left(\frac{30 + .5 - 20}{\sqrt{200(.1)(.9)}}\right) = \Phi(2.475).$$

Since $\Phi(2.47) = .9932$ and $\Phi(2.48) = .9934$, by interpolation, $\Phi(2.475) = .9933$.

7. To compare two different manufacturers of widget shafts, two samples of sizes n_1 and n_2 were ordered from the first and second manufacturers. Suppose X_1 of the first sample were nonconforming and X_2 of the second sample were nonconforming. Show that $\hat{\theta} = \frac{X_1}{n_1} - \frac{X_2}{n_2}$ is an unbiased estimator for $\theta = p_1 - p_2$, where p_i is the population proportion of nonconforming shafts from manufacturer i . Derive the formula for standard error of $\hat{\theta}$. What assumptions are you making in your derivation?

The statistic $\hat{\theta}$ is unbiased if $E(\hat{\theta}) = \theta$. We are given that $X_i \sim \text{Binomial}(n_i, p_i)$ are binomial variables. Hence $E(X_i) = p_i n_i$ and $V(X_i) = n_i p_i q_i$. Using the linearity of expectations

$$E(\hat{\theta}) = E\left(\frac{X_1}{n_1} - \frac{X_2}{n_2}\right) = \frac{E(X_1)}{n_1} - \frac{E(X_2)}{n_2} = \frac{n_1 p_1}{n_1} - \frac{n_2 p_2}{n_2} = p_1 - p_2.$$

The standard error is $\sigma_{\hat{\theta}}^2 = V(\hat{\theta})$. Let us assume that X_1 and X_2 are independent. Then the formula for variance of a linear combination yields

$$\sigma_{\hat{\theta}} = \sqrt{V\left(\frac{X_1}{n_1} - \frac{X_2}{n_2}\right)} = \sqrt{\frac{V(X_1)}{n_1^2} + \frac{V(X_2)}{n_2^2}} = \sqrt{\frac{n_1 p_1 q_1}{n_1^2} + \frac{n_2 p_2 q_2}{n_2^2}} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}.$$

8. In the study "Vitamin C Retention in Reconstituted Frozen Orange Juice," (VPI Department of Human Nutrition and Foods, 1972), levels of vitamin C were measured for twelve samples at two different time periods (3, 7 days) between when OJ concentrate was blended and when it was tested. Response is mg/l ascorbic acid. Test whether the data shows that there is a difference in mean vitamin levels at different times. State the assumptions. State the null and alternative hypotheses, test statistic, rejection region and your conclusion.

Sample	3 days	7 days	D
1	49.4	42.7	6.7
2	49.2	48.8	0.4
3	42.8	40.4	2.4
4	53.2	47.6	5.6
5	48.8	49.2	-0.4
6	44.0	44.0	0.0
7	44.0	42.0	2.0
8	42.4	43.2	-0.8
9	48.0	48.5	-0.5
10	47.0	43.3	3.7
11	48.2	45.2	3.0
12	49.6	47.6	2.0

Let X_i be the acid level for the i -th sample measured at three days and Y_i be the acid level of the same sample measured at seven days. X_i and Y_i are therefore not independent, and we must use the paired test. We have recorded the random variable $D_i = X_i - Y_i$ in the table. The sample is small ($n = 12 \leq 40$) and σ_D is unknown so we use a paired t -test. We assume that D_i is a random sample taken from a normal distribution $D_i \sim N(\mu_D, \sigma_D)$. The null and alternative hypotheses are

$$\begin{aligned} \mathcal{H}_0 &: \mu_D = 0; \\ \mathcal{H}_a &: \mu_D \neq 0. \end{aligned}$$

The test statistic is

$$T = \frac{\bar{D}}{S_D/\sqrt{n}}.$$

The null hypothesis is rejected at the α level of significance if $|t| \geq t_{\nu, \alpha/2}$ since it is a two-tailed test where $\nu = n - 1 = 11$ is the degrees of freedom. The calculator gives $\bar{D} = 2.008333333$ and $S_D = 2.440364478$. Hence

$$t = \frac{2.008}{2.440/\sqrt{12}} = 2.851.$$

Thus, interpolating Table A.8,

$$\frac{1}{2}P = P(T > 2.851) = .008.$$

Thus the P -value is .016. Similarly we could have used Table A.5: $t_{11, .010} = 2.718$ and $t_{11, .005} = 3.106$ so interpolating, $t_{11, .008} = 2.851$. Thus the mean ascorbic acid level at the two times is significantly different.

9. The Wanship Widgets Company wishes to test whether the average tensile strength of steel wire exceeds the tensile strength of iron wire by at least 10 kilograms. To test the claim, 50 pieces of each type of wire are tested under similar conditions. The steel wire had an average tensile strength of 88.9 kilograms with a standard deviation of 6.28 kilograms, while the iron wire had an average strength of 76.9 kilograms with a standard deviation of 5.61 kilograms. State the null and alternative hypotheses, the test statistic, and the rejection region. Test at the .05 level of significance and state your conclusion. Why is your statistic valid? Compute the probability that the test will reject the null hypothesis assuming that the mean tensile strength of the steel wire is actually 14 kilograms higher than that of the iron. Assume the same standard deviations.

Let X_i be the tensile strength of the steel wire and Y_j the tensile strength of the iron wire. We are assuming that all the X_i 's and Y_j 's are independent, that X_i comes from a distribution with population mean μ_1 and population standard deviation σ_1 and Y_j comes from a distribution with population mean μ_2 and population standard deviation σ_2 . Since both populations are large, $n_1 = n_2 = 50 > 40$, we may use the two sample z -test, even though the σ_i 's are unknown. We also cannot assume that $\sigma_1 = \sigma_2$ so we must not pool the variance computation. The null and alternative hypotheses are

$$\begin{aligned}\mathcal{H}_0 : \mu_1 - \mu_2 &= \Delta_0 = 10; \\ \mathcal{H}_a : \mu_1 - \mu_2 &> 10.\end{aligned}$$

The test statistic is

$$Z = \frac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}.$$

The null hypothesis is rejected at the α level of significance if $z \geq z_\alpha$ since it is a one-tailed test. We have for $\alpha = .05$, $z_{.05} = 1.645$. Computing,

$$\frac{88.9 - 76.9 - 10.0}{\sqrt{\frac{6.28^2}{50} + \frac{76.9^2}{50}}} = 1.679.$$

Thus $1.679 > 1.645$ so the null hypothesis is rejected: there is strong evidence that the mean tensile strength of steel wire exceeds the mean tensile strength of iron wire by more than 10 kilograms.

If the actual difference is $\Delta' = 14$ kilograms, then the probability of accepting the null hypothesis is

$$\begin{aligned}\beta(\Delta') &= P(Z < z_\alpha \mid \mu_1 - \mu_2 = \Delta') \\ &= P\left(\bar{X} - \bar{Y} \leq \Delta_0 + z_\alpha \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \mid \mu_1 - \mu_2 = \Delta'\right) \\ &= P\left(Z = \frac{\bar{X} - \bar{Y} - \Delta'}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \leq \frac{\Delta_0 - \Delta'}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} + z_\alpha \mid \mu_1 - \mu_2 = \Delta'\right) \\ &= \Phi\left(\frac{10 - 14}{\sqrt{\frac{6.28^2}{50} + \frac{76.9^2}{50}}} + 1.645\right) = \Phi(-1.714) = .0432.\end{aligned}$$

We have interpolated Table A.5: $\Phi(-1.71) = .0436$ and $\Phi(-1.72) = .0427$.