Sample Problems

1. In 2008, The Centers for Disease Control and Prevention (CDC) reported 11,413 individuals contracted HIV/AIDS via heterosexual contact. Of these, 4,301 were males and 7,112 were females.\(^1\) Is there evidence that females are more likely to contract HIV/AIDS via heterosexual contact than males?

2. A clinical trial was conducted to assess the effectiveness of a new method for treating early-stage melanoma patients who have had cancerous lymph nodes removed. The conventional treatment is often called watch-and-wait after surgery and the new, more aggressive treatment (LM/SNB) involves regular biopsies of the remaining lymph nodes. After 5 years, 21 of 30 subjects in the LM/SNB treatment group (group 1) were in good health whereas only 15 of 28 subjects in the watch-and-wait treatment group (group 2) were in good health. The proponents of the LM/SNB treatment ask you to carry out a hypothesis test.

3. Is resistance exercise as effective as aerobic exercise for body-weight management? Melanson et. al (2002)\(^2\) compared 24-hour energy expenditure and macronutrient oxidation elicited by comparable bouts of stationary cycling and weightlifting. 24-hour energy expenditure was measured in 10 non-obese male subjects on three occasions using whole-room indirect calorimetry. Stationary cycling and weightlifting were compared with a non-exercise control day. During stationary cycling, subjects exercised for 49 minutes on average at 70% of VO\(_2\) max and expended an average of 546 kcal. During weightlifting, subjects performed a 70-min circuit consisting of four sets of 10 different exercises at 70% of exercise-specific 1-repetition maximum and expended on average 448 kcal.

The sample means 24-hour of energy expenditure on biking days and and weightlifting days were \(\bar{y}_1 = 2787\) kcal and \(\bar{y}_2 = 2739\) kcal, respectively, and the sample standard deviations were \(s_1 = 76\) kcal and \(s_2 = 106\) kcal. For the control group, \(\bar{y}_3 = 2260\) kcal and \(s_3 = 96\) kcal.

The researchers reported that in men, resistance exercise has a similar effect on 24-h EE and macronutrient oxidation as a comparable bout of aerobic exercise. Neither exercise produced an increase in 24-h fat oxidation above that observed on a nonexercise control day. Confirm, by means of 95% confidence interval that the researchers have presented a statistically defensible conclusion.

---

\(^1\)http://www.cdc.gov/hiv/topics/surveillance/basic.htm#hivaidsexposure

4. In a 1990 article\textsuperscript{3}, researchers reported the results of a study that controlled for genetic and socioeconomic differences by examining 15 pairs of monozygotic twins, where one of the twins was schizophrenic and the other was not. The researchers used magnetic resonance imaging to measure the volumes (in cm\textsuperscript{3}) of the left hippocampus of the twins brains. Data are summarized below.

\begin{tabular}{|c|c|}
\hline
Unaffected & Affected \\
\hline
$\bar{y}$ & 1.758 & 1.560 \\
$s$ & .2424 & .3012 \\
\hline
\end{tabular}

The sample mean difference was $d = .199$ cm\textsuperscript{3} and the sample standard deviation was $sd = .238$ cm\textsuperscript{3}. Is there evidence of a difference in mean volume of the left hippocampus between schizophrenic and healthy individuals?

5. The following data are metabolic expenditures for 8 patients admitted to a hospital for reasons besides trauma and for 7 admitted for multiple fractures (trauma).\textsuperscript{4} It is believed that trauma sometimes results in catabolism, a destructive metabolic wasting process in which proteins and fat are broken down with the release of energy.

\begin{tabular}{|c|c|c|c|c|c|c|c|}
\hline
Metabolic expenditure (kcal/kg/day) & & & & & & & \\
Non-trauma patients & 20.1 & 22.9 & 18.8 & 20.9 & 22.7 & 21.4 & 20.0 & 20.9 \\
Trauma patients & 38.5 & 25.8 & 22.0 & 23.0 & 37.6 & 30.0 & 24.5 \\
\hline
\end{tabular}

Assess the evidence supporting the contention that trauma patients experienced greater metabolic expenditure, on average, than non-trauma patients.


Solutions

1. HIV problem. Either a confidence interval or a hypothesis test can be used to answer the question; a hypothesis test is more definitive and so preferable. If females are as likely as males to contract HIV/AIDS via heterosexual contact, then the proportion of females that contract HIV/AIDS via heterosexual contact would be \( p = .5 \). A test of this hypothesis is appropriate. Let \( p \) denote the proportion of females among all individuals that contract HIV/AIDS via heterosexual contact. The hypotheses of interest are

\[
H_0 : p = .5 \quad \text{and} \quad H_a : p > .5.
\]

The estimate of \( p \) is \( \hat{p} = 7112/11413 = .623 \). Note that \( p_0 = q_0 = .5 \). The test statistic is

\[
Z = \frac{\hat{p} - p_0}{\sqrt{p_0q_0/n}} = \frac{.623 - .5}{\sqrt{.5 \times .5 / 7112}} = \frac{.123}{.00468} = 26.3.
\]

The p-value is \( P(Z > 26.3|H_0) < .0000001 \) so there is very strong evidence that females are more likely than males to contract HIV/AIDS via heterosexual contact.

A discussion of the necessary conditions is needed. The data are cases reported to the CDC. Every doctor has a legal and professional responsibility to report cases so I’m willing to view the 7112 reported cases for 2008 as a representative sample of a larger population of cases. The 10% and success/failure conditions are easily satisfied.


Step 1 identifies to which of the situation the key. The problem is a comparison of population proportions where the populations are

(a) Early-stage melanoma patients who have had cancerous lymph nodes removed and receive the LM/SNB treatment. \( p_1 \) is the proportion of this population that are in good health 5 years after treatment.

(b) Early-stage melanoma patients who have had cancerous lymph nodes removed and receive the watch-and-wait after surgery treatment. \( p_2 \) is the proportion of this population that are in good health 5 years after treatment.

Step 2 sets up the hypotheses for testing whether the LM/SNB treatment is more effective than the watch-and-wait treatment.

\[
H_0 : p_1 - p_2 = 0
\]
Step 3 is identifies the test statistic and computes the terms necessary to evaluate the test statistic. The test statistic is the two-proportion $z$-statistic:

$$ Z = \frac{\hat{p}_1 - \hat{p}_2}{\hat{\sigma}_{pooled}(\hat{p}_1 - \hat{p}_2)} $$

The necessary terms are

(a) The sample proportion for each treatment group. $\hat{p}_1 = 21/30 = .7$ and $\hat{p}_2 = 15/28 = .536$.

(b) The pooled sample proportion $\hat{p}_{pooled} = \frac{21 + 15}{30 + 28} = .621$.

(c) The standard error of the difference in the two proportions.

$$ \hat{\sigma}_{pooled}(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_{pooled}\hat{q}_{pooled}}{n_1 + n_2}} $$

$$ = \sqrt{\frac{.621 \times .379}{30 + 28}} = .127. $$

Step 4 is to check the large-sample and independence conditions.

i. Independence: the problem is drawn from a real clinical trial funded by National Institutes of Health so I’m comfortable assuming the samples are independent.

ii. The success/failure condition is nearly satisfied because $n_1\hat{p}_1 = 21 > 10$, $n_1\hat{q}_1 = 9 < 10$, $n_2\hat{p}_2 = 15 > 10$, $n_2\hat{q}_2 = 13 > 10$.

iii. The 10% condition is satisfied since the population sizes are very large.

Step 5 computes the test statistic and the p-value:

$$ z = \frac{\hat{p}_1 - \hat{p}_2}{\hat{\sigma}_{pooled}(\hat{p}_1 - \hat{p}_2)} $$

$$ = \frac{.7 - .536}{.127} $$

$$ = 1.29. $$

StatCrunch reports an area of .098 to the right of 1.29, so p-value = $P(Z \geq 1.29) = .098$.

Step 6 states a conclusion drawn from the results: there is some evidence that the LM/SNB treatment is more effective than the watch-and-wait treatment. The difference $\ .7 - .536 = .164$ is sufficiently large to be practically significant, and so there is strong motivation to conduct a second trial aimed at strengthening the evidence supporting $H_a$. 

$H_a : p_1 - p_2 > 0.$
3. **Exercise problem.** Apparently, the same men were used for all three activities but the results are reported as if there were three independent samples. A confidence interval based on paired data is preferable, but not possible. Instead, a confidence interval will be computed as if the data consisted of two independent samples. A 100(1 - 2α)% confidence interval for \( \mu_1 - \mu_2 \) is

\[
\bar{y}_1 - \bar{y}_2 \pm t^* \hat{\sigma}(\bar{y}_1 - \bar{y}_2) = \bar{y}_1 - \bar{y}_2 \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}
\]

The sample statistics are \( \bar{y}_1 = 2787 \) kcal and \( \bar{y}_2 = 2739 \) kcal, respectively, for the biking and weightlifting days. The sample standard deviations were \( s_1 = 76 \) kcal and \( s_2 = 106 \) kcal. Sample sizes were \( n_1 = 10 \) and \( n_2 = 10 \). The critical t-value, based on \( df = n_1 - 1 = 9 \) degrees of freedom is \( t^* = 2.262 \). The confidence interval is

\[
\bar{y}_1 - \bar{y}_2 \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 2787 - 2739 \pm 2.262 \sqrt{\frac{76^2}{10} + \frac{106^2}{10}} = 48 \pm 2.262 \times 41.24 = [-45.28, 141.28].
\]

The confidence interval comfortably contains 0 and so there is no evidence in these data of a difference in mean 24-hour energy expenditure.

4. **Twins’ hippocampus volume.** Evidence is assessed by means of a paired t-test. Let \( \mu_D \) denote the mean difference in left hippocampus volume of twins where one is affected by schizophrenia and the other is not; specifically, \( \mu_D = \mu_{\text{unaffected}} - \mu_{\text{affected}} \). The hypotheses are \( H_0 : \mu_D = 0 \) versus \( H_a : \mu_D \neq 0 \). The test statistic is

\[
T = \frac{\bar{d} - \Delta_0}{s_{\bar{d}} / \sqrt{n}} = \frac{.1986 - 0}{.238 / \sqrt{15}} = 3.228.
\]

Also \( p\)-value = \( 2P(T_{14} \geq 3.228) = .00606 \). There is very strong evidence of anatomical differences between schizophrenic and healthy individuals. A 95% confidence interval for \( \mu_D \) is \([.066, .331]\).

5. **Metabolic expenditure** Let \( \mu_1 \) and \( \mu_1 \) denote the expected metabolic expenditure of trauma and non-trauma patients, respectively. The hypotheses of interest are

\[
H_0 : \mu_1 - \mu_2 = 0 \text{ and } H_a : \mu_1 - \mu_2 > 0.
\]
The test statistic is the two-sample $t$:

\[
T = \frac{\bar{y}_1 - \bar{y}_2}{\hat{\sigma}(\bar{y}_1 - \bar{y}_2)} = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}.
\]

The sample statistics are $\bar{y}_1 = 28.77$, $\bar{y}_2 = 20.96$, $s_1 = 6.835$ and $s_2 = 1.379$. Thus,

\[
T = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{28.77 - 20.96}{\sqrt{\frac{6.835^2}{7} + \frac{1.379^2}{8}}} = \frac{7.809}{\sqrt{6.674 + .2718}} = 2.970.
\]

The p-value is approximately $P(T_6 \geq 2.970) = .0104$, which is strong evidence that the difference in expected metabolic expenditures between trauma and non-trauma patients is greater than zero. It can be concluded that the metabolic expenditure of trauma patients is greater, on average than that of other patients.