1. An airline finds that 4% of the persons who make reservations on a certain flight do not show up for the flight. If the airline sells 170 tickets for a flight with only 166 seats, what is the probability that a seat will be available for every person holding a reservation and planning to fly? (Don’t forget to use continuous correction if you apply normal approximation.)

0.90

2. Let $X$ be a random variable and $X_1, \ldots, X_n$ a random sample of $X$. An unbiased estimator for the variance of $X$ is ($\bar{X}$ is the sample mean.)

$$\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

3. A survey was conducted to evaluate the effectiveness of a new flu vaccine that had been administered in a community. The vaccine was provided in a two-shot sequence over a period of 3 weeks to those wishing to avail themselves of it. Some people received the two-shot sequence, some appeared only for the first shot, and the others received neither.

A survey of 1000 local inhabitants in the following spring provided the information shown in the following table. Do the data present sufficient evidence to indicate a dependence between the two classifications – vaccine category and occurrence of nonoccurrence of flu? Use $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>Status</th>
<th>No Vaccine</th>
<th>One Shot</th>
<th>Two Shots</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flu</td>
<td>54</td>
<td>29</td>
<td>53</td>
<td>136</td>
</tr>
<tr>
<td>No Flu</td>
<td>259</td>
<td>80</td>
<td>525</td>
<td>864</td>
</tr>
<tr>
<td>Total</td>
<td>313</td>
<td>109</td>
<td>578</td>
<td>1000</td>
</tr>
</tbody>
</table>

28.9 > 5.99147, yes.

4. The numbers of accidents experienced by machinists were observed for a fixed period of time, with the results as shown in the following table. Test, at the 5% level of significance, the hypothesis that the data come
from a Poisson distribution.

<table>
<thead>
<tr>
<th>Accidents per Machinist</th>
<th>Frequency of Observation (number of machinists)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>296</td>
</tr>
<tr>
<td>1</td>
<td>74</td>
</tr>
<tr>
<td>2</td>
<td>26</td>
</tr>
<tr>
<td>3 or more</td>
<td>18</td>
</tr>
</tbody>
</table>

I have calculated the (sample) mean number of accidents per machinist for you: \( \bar{x} = 0.4831 \).

5. A study was conducted by the Florida Game and Fish Commission to assess the amounts of chemical residues found in the brain tissue of brown pelicans. In a test for DDT, random samples of \( n_1 = 11 \) juveniles and \( n_2 = 13 \) nestlings produced the results shown in the following table (measurements in parts per million, ppm).

<table>
<thead>
<tr>
<th></th>
<th>Juveniles</th>
<th>Nestlings</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_1 )</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>( \bar{y}_1 )</td>
<td>0.041</td>
<td>0.026</td>
</tr>
<tr>
<td>( s_1 )</td>
<td>.017</td>
<td>.006</td>
</tr>
</tbody>
</table>

Is there sufficient evidence, at the 5% significance level, to support concluding that the variance in measurements of DDT levels is greater for juveniles than it is for nestlings? \( 8.027 > 2.75, \text{yes.} \)

6. Independent random samples of weekend and weekday shoppers were selected and the amount spent per trip to the mall was recorded as shown in the following table:

<table>
<thead>
<tr>
<th></th>
<th>Weekends</th>
<th>Weekdays</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_1 )</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>( \bar{y}_1 )</td>
<td>$78</td>
<td>$67</td>
</tr>
<tr>
<td>( s_1 )</td>
<td>$22</td>
<td>$20</td>
</tr>
</tbody>
</table>

Is there sufficient evidence to claim that there is a difference in the average amount spent per trip on weekends and weekdays? Use \( \alpha = 0.05 \).

\( 1.4329 < 2.048, \text{no.} \)
7. A chemical process has produced, on the average, 800 tons of chemical per day. The daily yields for the past week are 775, 800, 780, 790, and 788 tons. Do these data indicate that the average yield is less than 800 tons and hence that something is wrong with the process? Test at 5% level of significance. (I have calculated $\bar{x} = 786.6$, and $s = 9.633$.)

$-3.11 < -2.132$, yes.

8. Calculate the sample mean $\bar{x}$ and sample standard deviation $s$ for the following data: 1, 4, 9, 16, 25.

$\bar{x} = 11$, $s = 9.6695$.

9. A check-cashing service found that approximately 5% of all checks submitted to the service were bad. After instituting a check-verification system to reduce its losses, the service found that only 45 checks were bad in a random sample of 1134 that were cashed. Does sufficient evidence exist to affirm that the check-verification system reduced the proportion of bad checks? Use $\alpha = 6.06\%$.

$-1.594 < -1.55$, yes.

10. A check-cashing service found that approximately 5% of all checks submitted to the service were bad. After instituting a check-verification system to reduce its losses, the service found that only 45 checks were bad in a random sample of 1200 that were cashed. We wish to test whether the check-verification system reduced the proportion of bad checks? What attained significance level (i.e. p-value) is associated with the test? Is the hypothesis $p < 0.05$ accepted at 3% significance level?

0.0233, yes.

11. Two brands of refrigerators, denoted A and B, are each guaranteed for 1 year. In a random sample of 50 refrigerators of brand A, 12 were observed to fail before the guarantee period ended. An independent random sample of 60 brand B refrigerators also revealed 12 failures during the guarantee period. Estimate the true difference $(p_1 - p_2)$ between proportions of failures during the guarantee period, with confidence coefficient approximately .98.

$[-.1451, .2251]$.

12. An experimenter wanted to check the variability of measurements obtained by using equipment designed to measure the volume of an audio source. Three independent measurements recorded by this equipment
for the same sound were 4.1, 5.2, and 10.2. Estimate $\sigma^2$ with confidence coefficient .90.

\[(3.53, 205.24)\]

13. In a linear model $Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \epsilon$, suppose we have the following data

\[
\begin{array}{ccc}
x_1 & x_2 & y \\
2.2 & 1.2 & 3.1 \\
0.3 & 1.4 & 1.5 \\
5 & 2.1 & 4.8 \\
1.1 & 2.6 & 3.8 \\
3.3 & 1.1 & 1.6
\end{array}
\]

Let $X$ be the matrix

\[
\begin{bmatrix}
1 & 2.2 & 1.2 \\
1 & 0.3 & 1.4 \\
1 & 5 & 2.1 \\
1 & 1.1 & 2.6 \\
1 & 3.3 & 1.1
\end{bmatrix}
\]

and $Y$ be the vector

\[
\begin{bmatrix}
3.1 \\
1.5 \\
4.8 \\
3.8 \\
1.6
\end{bmatrix}
\]

Then

\[
(X'X)^{-1} = \begin{bmatrix}
2.2853 & -0.1694 & -1.0013 \\
-0.1694 & 0.0730 & -0.0025 \\
-1.0013 & -0.0025 & 0.5996
\end{bmatrix},
\]

\[
X'Y = \begin{bmatrix}
14.8 \\
40.73 \\
27.54
\end{bmatrix},
\]

and

\[
(X'X)^{-1}X'Y = \begin{bmatrix}
-0.6521 \\
0.3949 \\
1.5906
\end{bmatrix}.
\]

Using this information, find $\hat{\beta}_1$, the least-squares unbiased estimator of $\beta_1$.

\[0.3949\]
14. (continued) Assume that $V(\epsilon) = \sigma^2$, a constant independent of the variables $x_1, x_2$. Find the least-squares unbiased estimator of $\sigma^2$.  
0.8305

15. (optional, continued) Under the hypothesis of normality, is there sufficient evidence to conclude that $\beta_1 > 0$ as oppose to $\beta_1 = 0$ at 5% level of significance? (Please show you work on the blank space below or the back.)  
1.60353;2.920, no