Instructions.

1. Attempt all questions.

2. Show all the steps of your work clearly. Correct final answers without enough work may receive no credit.

3. Find the needed tables at the end of the packet. You may tear this sheet free.

4. Don’t worry if you don’t finish the test. Just try to score as many points as you can.

5. Good luck!

<table>
<thead>
<tr>
<th>Question</th>
<th>Points</th>
<th>Your Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Q2</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Q3</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Q4</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Q5</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Q6</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Q7</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>90</td>
<td></td>
</tr>
</tbody>
</table>
A team of scientists developed a flesh-regenerating virus known as the T-virus. Unfortunately, the virus caused an apocalyptic zombie outbreak. A cure was later developed, and its success probability was \( p = 0.73 \). Suppose \( n = 16 \) zombies were captured and injected with the cure. Let \( Y \) be the number of cured zombies. Using a normal approximation with continuity correction and taking \( \hat{p} = Y/n \), evaluate

\[
P(0.8 > \hat{p} \geq 0.6).
\]

Solution: Multiply both sides by \( n \), apply the continuity correction, apply the CLT:

\[
P(0.8 > \hat{p} \geq 0.6) = P(12.8 > Y \geq 9.6)
\]

\[
= P(12.3 > Y \geq 9.1)
\]

\[
= P\left(\frac{12.3 - np}{\sqrt{np(1-p)}} > Z \geq \frac{9.1 - np}{np(1-p)}\right)
\]

\[
= P(0.35 > Z \geq -1.45)
\]

\[
= 0.5633014,
\]

noting that \( np = 11.68 \) and \( \sqrt{np(1-p)} = 1.775838 \).

Can we trust this approximation? Why or why not?

Solution: No, because \( n(1-p) = 4.32 < 5 \).
Q2...[10 points] The mean lifespan of a certain species of rat is studied. Scientists need the error margin of a 95% confidence interval for $\mu$ not to exceed 50 days. A pilot study gives $\sigma^2 = 10,000$. Find the smallest whole number $n$ which satisfies this requirement.

Solution: the error margin takes the form $2\sigma/\sqrt{n}$ (NOTE, the error margin is half the length of the CI), which we need to be less than 50. Thus solve

$$2\sigma/\sqrt{n} \leq 50$$
$$\Rightarrow 4\sigma^2/n \leq 2500$$
$$\Rightarrow 40000/n \leq 2500$$
$$\Rightarrow 40000/2500 \leq n$$
$$\Rightarrow 16 \leq n$$

Giving a final small whole number sample size of 16.
A chemical process in an industrial plant after 12 iterations resulted in a mean yield of 12.2 kg (SD = 2.1 kg). Engineers added a new catalyst, repeated the process 8 times, and recorded a mean yield of 14.3 kg (SD = 1.9 kg). Using the pooled standard error, the liberal degrees of freedom method (df = \( n_1 + n_2 - 2 \)), and \( \alpha = .05 \), test

\[ H_0 : \mu_1 - \mu_2 = 0 \quad \text{versus} \quad H_A : \mu_1 - \mu_2 \neq 0 \]

Solution: Find \( s_{pooled}^2 \) first:

\[
s_{pooled}^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = 4.099.
\]

Second, calculate the pooled standard error

\[
SE_{pooled} = \sqrt{s_{pooled}^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)} = \sqrt{4.099 \left( \frac{1}{12} + \frac{1}{8} \right)} = 0.924.
\]

Third, calculate the observed value of the test statistic

\[
t = \frac{\bar{y}_1 - \bar{y}_2}{SE_{pooled}} = \frac{-2.1}{0.924} = -2.27
\]

We reject the null hypothesis if

\[ 2.27 = |t| > t_{\alpha/2, n_1 + n_2 - 2} = t_{.025,18} = 2.1009. \]

So we reject \( H_0 \) at \( \alpha = .05 \).
Q4]...[10 points] Fill in the following table (write you answer after (a) and (b) below).

<table>
<thead>
<tr>
<th>$H_0$ false</th>
<th>$H_0$ true</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retain $H_0$</td>
<td>(a) Correct</td>
</tr>
<tr>
<td>Reject $H_0$</td>
<td>Correct (b)</td>
</tr>
</tbody>
</table>

(a) Type II Error  
(b) Type I Error

Now, write a short paragraph (only two sentences are needed) which interprets the two types of error in a hypothesis test in the context of the court case metaphor used in class.

A type I error occurs when an innocent person is sent to jail. A type II error occurs when a guilty person is exonerated.
Q5][10 points] In a 2-sample $t$-test with $\alpha = .05$, $H_A : \mu_1 \neq \mu_2$, and df= 12, the observed value of the test statistic equaled $t = -2.05$. Find (1) the critical value, (2) whether we reject $H_0 : \mu_1 = \mu_2$ in this case, and (3) whether we reject the null hypothesis if $H_A : \mu_1 < \mu_2$ instead.

Solution.

1. The critical value in this case is $t_{\alpha/2,df} = t_{.025,12} = 2.1788$.

2. Since $|t| = 2.05 < 2.1788 = t_{.025,12}$, we retain $H_0$.

3. The critical value in this case is $-t_{\alpha,df} = -t_{.05,12} = -1.7823$. Since $t = -2.05 < -1.7823$, we reject $H_0$. 
Scientists study the resting metabolic rate (RMR), or the number of calories burned per hour at rest, of a human male that weighs 71kg. It is known that, while RMR is linked to body weight, the variability of RMR is not and is \( \sigma = 4.8 \) calories per hour. The subject in question undergoes 21 RMR measurements, giving \( s = 5.1 \) and \( \bar{y} = 72.3 \) calories per hour. The measurements follow a normal distribution. Calculate the shortest 95% confidence interval for \( \mu \), the true mean RMR for this subject, using the available information. Justify your decision.

Solution. Since the data are normal and \( \sigma \) is known, we use the following CI formula

\[
\bar{y} \pm z_{0.025} \sigma / \sqrt{n}.
\]

Plugging in the numbers, we obtain

\[
72.3 \pm 1.96 \times 4.8 / \sqrt{21} = (70.25, 74.36).
\]
Q7]... [15 points] Researchers believe that a certain synthetic androgen (Drug A) will increase free throw accuracy of college basketball players. A player with a known free throw success probability of \( p = 0.93 \) is injected with the chemical. The player misses 3 free throws in 1000 independent attempts. Write down a mathematical expression using the binomial formula (with the correct numbers plugged in) which will evaluate to the p-value under the following hypotheses:

\[
H_0 : p = 0.93 \quad \text{versus} \quad H_A : p > 0.93.
\]

(Hint: The solution follows the steps of the street magician example.)

Solution: The p-value is the probability of viewing data as or more extreme than the observed data under the assumption that \( p = 0.93 \). Thus, the p-value is equal to

\[
\binom{1000}{997} (0.93)^{997} (0.07)^3 + \binom{1000}{998} (0.93)^{998} (0.07)^2 + \binom{1000}{999} (0.93)^{999} (0.07) + \binom{1000}{1000} (0.93)^{1000}
\]