1. If $X_1, X_2, \ldots, X_n$ is a random sample from an infinite Bernoulli population with parameter θ, then $\overline{X}$ is the proportion of successes in n trials.

   (a) **[4 pts.]** Verify that $E(\overline{X}) = \theta$.

$$\begin{aligned} E(X_i) &= \theta, \text{ for } i = 1, 2, \ldots, n \\ E(\overline{X}) &= E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) \\ &= \frac{1}{n}\sum_{i=1}^{n}\theta = \frac{1}{n}(n\theta) = \theta \end{aligned}$$

+1 for writing $\overline{X}$ in terms of X_i's
+1 for correct distribution of expected value
+1 for using $\mu = \theta$ for Bernoulli
+1 for correct final answer

   (b) **[4 pts.]** Verify that $var(\overline{X}) = \frac{\theta(1-\theta)}{n}$.

$$\begin{aligned} var(X_i) &= \theta(1-\theta), \text{ for } i = 1, 2, \ldots, n \\ var(\overline{X}) &= var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} var(X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\theta(1-\theta) \\ &= \frac{1}{n^2}\left[n\theta(1-\theta)\right] = \frac{\theta(1-\theta)}{n} \end{aligned}$$

+1 for writing $\overline{X}$ in terms of X_i's
+1 for correct distribution of variance
+1 for using $\sigma^2 = \theta(1-\theta)$ for Bernoulli
+1 for correct final answer

2. Suppose $X_1, X_2, \ldots, X_{n_1}$ is a random sample from an infinite Bernoulli population with parameter θ_1 and $Y_1, Y_2, \ldots, Y_{n_2}$ is a random sample from an infinite Bernoulli population with parameter θ_2. Then, $\overline{X}$ is the proportion of successes in n_1 trials for the first sample and $\overline{Y}$ is the proportion of successes in n_2 for the second sample. It is assumed that the X's and the Y's are independent.

   (a) **[4 pts.]** Verify that $E(\overline{X} - \overline{Y}) = \theta_1 - \theta_2$.

From exercise 1 (a) above,
$E(\overline{X}) = \theta_1$ and $E(\overline{Y}) = \theta_2$.

$$E(\overline{X} - \overline{Y}) = E(\overline{X}) - E(\overline{Y}) = \theta_1 - \theta_2$$

+1 for using exercise 1(a) to get $E(\overline{X}) = \theta_1$
+1 for using exercise 1(a) to get $E(\overline{Y}) = \theta_2$
+1 for correct distribution of expected value
+1 for correct final answer

(b) **[4 pts.]** Verify that $var(\overline{X}-\overline{Y}) = \frac{\theta_1(1-\theta_1)}{n_1} + \frac{\theta_2(1-\theta_2)}{n_2}$.

From exercise 1 (b) above, $var(\overline{X}) = \frac{\theta_1(1-\theta_1)}{n_1}$ and $var(\overline{Y}) = \frac{\theta_2(1-\theta_2)}{n_2}$.

$$\begin{aligned} var(\overline{X}-\overline{Y}) &= var(\overline{X}) + (-1)^2 var(\overline{Y}) \\ &= \frac{\theta_1(1-\theta_1)}{n_1} + \frac{\theta_2(1-\theta_2)}{n_2} \end{aligned}$$

+1 for using exercise 1(b) to get $var(\overline{X}) = \frac{\theta_1(1-\theta_1)}{n_1}$

+1 for using exercise 1(b) to get $var(\overline{Y}) = \frac{\theta_2(1-\theta_2)}{n_2}$

+1 for correct distribution of variance

+1 for correct final answer

3. For random samples from an infinite population, what happens to the standard error of the mean (state whether it increases or decreases and find the factor by which it is multiplied or divided) if the sample size is

(a) **[4 pts.]** increased from 30 to 120;

$$\begin{aligned} \frac{\sigma}{\sqrt{30}} \times m &= \frac{\sigma}{\sqrt{120}} \\ m &= \frac{\sigma}{\sqrt{120}} \times \frac{\sqrt{30}}{\sigma} = \sqrt{\frac{30}{120}} = \frac{1}{\sqrt{4}} = \frac{1}{2} \end{aligned}$$

The standard error DECREASES and it is multiplied by $\frac{1}{2}$ (or divided by 2)

+1 for showing some work

+2 for "decrease"

+1 for divided by 2

(b) **[4 pts.]** decreased from 450 to 50.

$$\begin{aligned} \frac{\sigma}{\sqrt{450}} \times m &= \frac{\sigma}{\sqrt{50}} \\ m &= \frac{\sigma}{\sqrt{50}} \times \frac{\sqrt{450}}{\sigma} = \sqrt{\frac{450}{50}} = \sqrt{9} = 3 \end{aligned}$$

The standard error INCREASES and it is multiplied by 3

+1 for showing some work

+2 for "increase"

+1 for multiply by 3

4. A random sample of size $n = 100$ is taken from an infinite population with mean $\mu = 75$ and variance $\sigma^2 = 256$. With what probability can we assert that the value of $\overline{X}$ will fall between 67 and 83

(a) **[4 pts.]** using the Weak Law of Large Numbers (Theorem 8.2);

The Weak Law of Large Numbers says $P(\mu - c < \overline{X} < \mu + c) \geq 1 - \frac{\sigma^2}{nc^2}$.

$$\begin{aligned}
\mu - c &= 67 \implies c = 75 - 67 \\
\mu + c &= 83 \implies c = 83 - 75 \\
&\implies c = 8 \\
P(67 < \overline{X} < 83) &\geq 1 - \frac{256}{100(8^2)} \\
&= 1 - 0.04 = 0.96
\end{aligned}$$

+1 for using the WLLN
+1 for finding the correct value of c
+1 for correctly plugging in the values from the problem
+1 for correct final answer

(b) **[6 pts.]** using the Central Limit Theorem (Theorem 8.3) and the fact that for the standard normal $F(5) = 0.9999997$.

The Central Limit Theorem says that for large n, $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$$\begin{aligned}
& P(67 < \overline{X} < 83) \\
= \ & P\left(\frac{67-\mu}{\sigma/\sqrt{n}} < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{83-\mu}{\sigma/\sqrt{n}}\right) \\
= \ & P\left(\frac{67-75}{16/\sqrt{100}} < \frac{\overline{X}-75}{16/\sqrt{100}} < \frac{83-75}{16/\sqrt{100}}\right) \\
= \ & P(-5 < Z < 5) = F(5) - F(-5) \\
= \ & 2F(5) - 1 = 2(0.9999997) - 1 \\
= \ & 0.9999994
\end{aligned}$$

+1 for using the CLT
+1 for finding the correct z-score for the lower endpoint
+1 for finding the correct z-score for the lower endpoint
+1 for using symmetry of normal
+1 for using $F(5)$
+1 for correct final answer

5. For random samples of size n from a normal population with variance σ^2, use Theorem 8.11 to verify that the sampling distribution of S^2 has

According to Theorem 8.11, $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ which means that $E(\frac{(n-1)S^2}{\sigma^2}) = n-1$ and $var(\frac{(n-1)S^2}{\sigma^2}) = 2(n-1)$

(a) **[4 pts.]** $E(S^2) = \sigma^2$;

$$
\begin{aligned}
E(S^2) &= E\left(\frac{\sigma^2}{(n-1)}\frac{(n-1)S^2}{\sigma^2}\right) \\
&= \frac{\sigma^2}{(n-1)}E\left(\frac{(n-1)S^2}{\sigma^2}\right) \\
&= \frac{\sigma^2}{(n-1)}(n-1) = \sigma^2
\end{aligned}
$$

+1 for transforming to a known chi-square RV

+1 for correctly distributing the expected value

+1 for using the mean of the chi-square RV

+1 for correct final answer

(b) **[4 pts.]** $var(S^2) = \frac{2\sigma^4}{n-1}$.

$$
\begin{aligned}
var(S^2) &= var\left(\frac{\sigma^2}{(n-1)}\frac{(n-1)S^2}{\sigma^2}\right) \\
&= \left(\frac{\sigma^2}{(n-1)}\right)^2 var\left(\frac{(n-1)S^2}{\sigma^2}\right) \\
&= \left(\frac{\sigma^4}{(n-1)^2}\right)2(n-1) = \frac{2\sigma^4}{n-1}
\end{aligned}
$$

+1 for transforming to a known chi-square RV

+1 for correctly distributing the variance

+1 for using the variance of the chi-square RV

+1 for correct final answer

6. **[4 pts.]** Suppose that F is a random variable with F distribution and degrees of freedom ν_1, ν_2. Use the fact below to show that $E(F) = \frac{\nu_2}{\nu_2-2}$ for $\nu_2 > 2$.
FACT: If V is a chi-square random variable with ν_2 degrees of freedom, then $E(\frac{1}{V}) = \frac{1}{\nu_2-2}$.

$$
\begin{aligned}
E(F) &= E\left(\frac{U/\nu_1}{V/\nu_2}\right) = \frac{\nu_2}{\nu_1}E\left(\frac{U}{V}\right) \\
&= \frac{\nu_2}{\nu_1}E(U)E\left(\frac{1}{V}\right) = \frac{\nu_2}{\nu_1}(\nu_1)\left(\frac{1}{\nu_2-2}\right) \\
&= \frac{1}{\nu_2-2}
\end{aligned}
$$

This is only valid for $\nu_2 > 2$ or else the expected value would be negative, but we know the support for the F distribution is $f > 0$.

+1 for using definition of F from Theorem 8.14

+1 for correctly distributing the expected value

+1 for correctly using the FACT

+1 for correct final answer

7. **[4 pts.]** Verify that if T has a t distribution with ν degrees of freedom, then $X = T^2$ has an F distribution with $\nu_1 = 1$ and $\nu_2 = \nu$ degrees of freedom.

According to Theorem 8.12, then T can be written as

$$T = \frac{Z}{\sqrt{Y/\nu}}$$

where $Z \sim N(0,1)$, $Y \sim \chi^2_\nu$, and Z and Y are independent. Now, consider

$$X = T^2 = \frac{Z^2}{Y/\nu}$$

where $Z^2 \sim \chi^2_1$ (according to Theorem 8.7), $Y \sim \chi^2_\nu$, and Z^2 and Y are independent.
According to Theorem 8.14, then X has F distribution with $\nu_1 = 1$ and $\nu_2 = \nu$ degrees of freedom.

+1 for using the form of T from Theorem 8.12

+1 for specifying the distributions of Z and Y

+1 for transforming Z^2 and specifying it is chi-square with 1 DOF

+1 for using Theorem 8.14 to get the result

8. **[4 pts.]** A random sample of size $n = 225$ is to be taken from an exponential population with $\theta = 4$. Set up the Central Limit Theorem (Theorem 8.3) result using proper notation to determine probability that the mean of the sample will exceed 4.5, but DO NOT solve.

For an exponential random variable with $\theta = 4$, then $\mu = \theta = 4$ and $\sigma^2 = \theta^2 = 16$. Now, the Central Limit Theorem say that for large n, $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$.

$$\begin{aligned} P(\overline{X} > 4.5) &= P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} > \frac{4.5-\mu}{\sigma/\sqrt{n}}\right) \\ &= P\left(\frac{\overline{X}-4}{4/\sqrt{225}} > \frac{4.5-4}{4/\sqrt{225}}\right) \\ &= P(Z > 1.875) = 1 - F(1.875) \end{aligned}$$

where F is the CDF of a standard normal.

+1 for finding the mean and variance of the exponential RV

+1 for using the CLT to transform the probability to a known standard normal RV

+1 for correctly using the values from the problem to find the z-score for the endpoint

+1 for correct final answer

9. **[4 pts.]** A random sample of size $n = 100$ is to be taken from a normal population with $\sigma = 25$. What is the probability that the mean of the sample will differ from the mean of the population by 3 or more in either direction? Set up the solution using proper notation, but DO NOT solve.

According to Theorem 8.4, $\overline{X} \sim N(\mu, \sigma^2)$ for any n, which means that $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$.

$$\begin{aligned} & 1 - P(\mu - 3 < \overline{X} < \mu + 3) \\ = \; & 1 - P(-3 < \overline{X} - \mu < 3) \\ = \; & 1 - P\left(\frac{-3}{25/\sqrt{100}} < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{3}{25/\sqrt{100}}\right) \\ = \; & 1 - P(-1.2 < Z < 1.2) \\ = \; & 1 - [F(1.2) - F(-1.2)] \end{aligned}$$

where F is the CDF of a standard normal.

+1 for setting up the correct probability for $\overline{X}$

+1 for using Theorem 8.4 to get a known standard normal RV

+1 for using the values of the problem to get the endpoints

+1 for correct final answer

10. **[4 pts.]** The claim that the variance of a normal population is $\sigma^2 = 25$ is to be rejected if the sample variance of a random sample of size $n = 16$ exceeds 54.668 or is less than 12.102. What is the probability that this claim will be rejected even though the population variance really is $\sigma^2 = 25$? Set up the solution using proper notation, but DO NOT solve.

According to Theorem 8.11, $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$.

$$\begin{aligned} & P(S^2 > 54.668 \text{ or } S^2 < 12.102) \\ = \; & 1 - P(12.102 < S^2 < 54.668) \\ = \; & 1 - P\left(\tfrac{(n-1)12.102}{\sigma^2} < \tfrac{(n-1)S^2}{\sigma^2} < \tfrac{(n-1)54.668}{\sigma^2}\right) \\ = \; & 1 - P\left(\tfrac{(15)12.102}{25} < \tfrac{(n-1)S^2}{\sigma^2} < \tfrac{(15)54.668}{25}\right) \\ = \; & 1 - P(7.2612 < X^2 < 32.8008) \\ = \; & 1 - [F(32.8008) - F(7.2612)] \end{aligned}$$

where F is the CDF of a chi-square with 15 degrees of freedom.

+1 for setting up the correct probability for S^2

+1 for using Theorem 8.11 to get a known chi-square RV

+1 for using the values of the problem to get the endpoints

+1 for correct final answer (including DOF)