4.2.1. Let the observed value of the mean $\bar{X}$ and of the sample variance of a random sample of size 20 from a distribution that is $N(\mu, \sigma^2)$ be 81.2 and 26.5, respectively. Find respectively 90%, 95% and 99% confidence intervals for $\mu$. Note how the lengths of the confidence intervals increase as the confidence increases.

Solution. Since $n = 20$, $\bar{X} = 81.2$, $S^2 = 26.5$, we are looking for the confidence interval

$$[\bar{X} - \frac{S}{\sqrt{n}}t_{\alpha/2,n-1}, \bar{X} + \frac{S}{\sqrt{n}}t_{\alpha/2,n-1}].$$

If $100(1 - \alpha) = 90$, $\alpha = 0.1$, $t_{0.05,19} = 1.729$, then

$$[81.2 - \sqrt{\frac{26.5}{20}} \cdot 1.729, 81.2 + \sqrt{\frac{26.5}{20}} \cdot 1.729] = [79.2, 83.2].$$

If $100(1 - \alpha) = 95$, $\alpha = 0.05$, $t_{0.025,19} = 2.093$, then

$$[81.2 - \sqrt{\frac{26.5}{20}} \cdot 2.093, 81.2 + \sqrt{\frac{26.5}{20}} \cdot 2.093] = [78.8, 83.6].$$

If $100(1 - \alpha) = 99$, $\alpha = 0.01$, $t_{0.005,19} = 2.861$, then

$$[81.2 - \sqrt{\frac{26.5}{20}} \cdot 2.861, 81.2 + \sqrt{\frac{26.5}{20}} \cdot 2.861] = [77.9, 84.5].$$

4.2.12. Let $Y$ be $b(300, p)$. If the observed value of $Y$ is $y = 75$, find an approximate 90% confidence interval for $p$.

Solution. If $100(1 - \alpha) = 90$, $\alpha = 0.1$, $z_{\alpha/2} = 1.64$, and $k = 75$, $n = 300$, then compute

$$\left[\frac{k}{n} - z_{\alpha/2}\sqrt{\frac{k}{n}(1 - k/n)} \cdot \frac{k}{n} + z_{\alpha/2}\sqrt{\frac{k}{n}(1 - k/n)}\right].$$
we get

\[
\left[\frac{1}{4} - 1.64 \cdot \sqrt{\frac{1}{4} \cdot \frac{3}{4}} \cdot \frac{300}{300}, \frac{1}{4} + 1.64 \cdot \sqrt{\frac{1}{4} \cdot \frac{3}{4}} \cdot \frac{300}{300}\right] = [0.209, 0.291].
\]

4.2.17. It is known that a random variable \(X\) has a Poisson distribution with parameter \(\mu\). A sample of 200 observations from this distribution has a mean equal to 3.4. Construct an approximate 90\% confidence interval for \(\mu\).

**Solution.** If \(X\) has a Poisson distribution with parameter \(\mu\), then \(E(X) = \mu\) and \(Var(X) = \mu\). To construct a 90\% confidence interval for \(\mu\), we shall look at

\[
[\bar{X} - \frac{S}{\sqrt{n}} z_{\alpha/2}, \bar{X} + \frac{S}{\sqrt{n}} z_{\alpha/2}].
\]

Since for Poisson distribution, the expectation and variance is just the parameter \(\mu\), we can use the value of the sample mean for the sample variance \(S^2\), if we don’t have the information about \(S^2\). Hence the 90\% confidence interval for \(\mu\) is

\[
[3.4 - \sqrt{\frac{3.4}{200} \cdot 0.05}, 3.4 + \sqrt{\frac{3.4}{200} \cdot 0.05}] = [3.4 - \sqrt{\frac{3.4}{200} \cdot 1.64}, 3.4 + \sqrt{\frac{3.4}{200} \cdot 1.64}] = [3.186, 3.614].
\]

4.2.18. Let \(X_1, X_2, \cdots, X_n\) be a random sample from \(N(\mu, \sigma^2)\), where both parameters \(\mu\) and \(\sigma^2\) are unknown. A confidence interval for \(\sigma^2\) can be found as follows. We know that \((n-1)S^2/\sigma^2\) is a random variable with a \(\chi^2(n-1)\) distribution. Thus we can find constants \(a\) and \(b\) so that \(P((n-1)S^2/\sigma^2 < b) = 0.975\) and \(P(a < (n-1)S^2/\sigma^2 < b) = 0.95\).

(a) Show that this second probability statement can be written as

\[ P((n-1)S^2/b < \sigma^2 < (n-1)S^2/a) = 0.95. \]

(b) If \(n = 9\) and \(s^2 = 7.93\), find a 95\% confidence interval for \(\sigma^2\).

(c) If \(\mu\) is unknown, how would you modify the preceding procedure for finding a confidence interval for \(\sigma^2\)?

**Solution.** (a) \(a < (n-1)S^2/\sigma^2 \iff \sigma^2 < (n-1)S^2/a\);

and

\[ (n-1)S^2/\sigma^2 < b \iff \sigma^2 > (n-1)S^2/b. \]
Hence the second probability statement can be written as
\[ P((n - 1)S^2/b < \sigma^2 < (n - 1)S^2/a) = 0.95. \]

(b) Since \( \alpha = 0.05, n = 9, b = \chi_{\alpha/2, n-1}^2 = \chi_{0.025,8}^2 = 17.54 \), and \( a = \chi_{1-\alpha/2, n-1}^2 = \chi_{0.975,8}^2 = 2.18. \) Also \( s^2 = 7.93. \) Hence the 95% confidence interval is
\[ [(n - 1)S^2/b, (n - 1)S^2/a] = \left[ \frac{8 \cdot 7.93}{17.54}, \frac{8 \cdot 7.93}{2.18} \right] = [3.62, 29.10]. \]

(c) In this case, define \( \tilde{S}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu)^2, \) then \( \frac{n\tilde{S}^2}{\sigma^2} \sim \chi^2(n). \) We can follow similar procedure by finding \( a \) and \( b \) from \( P(n\tilde{S}^2/\sigma^2 < b) = 0.975 \) and \( P(a < n\tilde{S}^2/\sigma^2 < b) = 0.95. \) Then the new confidence interval should be \([n\tilde{S}^2/b, n\tilde{S}^2/a].\)

4.2.21. Let two independent random samples, each of size 10, from two normal distributions \( N(\mu_1, \sigma^2) \) and \( N(\mu_2, \sigma^2) \) yields \( \bar{x} = 4.8, s_1^2 = 8.64, \bar{y} = 5.6, s_2^2 = 7.88. \) Find a 95% confidence interval for \( \mu_1 - \mu_2.\)

**Solution.** We’ll use the formula to find the confidence interval
\[ ((\bar{x} - \bar{y}) - t_{\alpha/2, n_2-2}\sqrt{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, ((\bar{x} - \bar{y}) + t_{\alpha/2, n_2-2}\sqrt{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}) \]

With information we get
\[ s_p = \frac{9 \cdot 8.64 + 9 \cdot 7.88}{9 + 9} = 8.26, \]

So the 95% confidence interval for \( \mu_1 - \mu_2 \) is
\[ (4.8 - 5.6 - t_{0.025,18} \cdot 2.874 \cdot 0.447, 4.8 - 5.6 + t_{0.025,18} \cdot 2.874 \cdot 0.447) = (-3.499, 1.899) \]

4.2.22. Let two independent random variables, \( Y_1 \) and \( Y_2, \) with binomial distributions that have parameters \( n_1 = n_2 = 100, p_1, \) and \( p_2, \) respectively, be observed to be equal to \( y_1 = 50 \) and \( y_2 = 40. \) Determine an approximate 90% confidence interval for \( p_1 - p_2. \)

**Solution.** Let \( \hat{p}_1 = \frac{50}{100} = 0.5, \hat{p}_2 = \frac{40}{100} = 0.4, \) then an 95% confidence interval for \( p_1 - p_2 \) is
\[ (\hat{p}_1 - \hat{p}_2 - z_{0.05} \cdot \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}), \hat{p}_1 - \hat{p}_2 + z_{0.05} \cdot \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}) \]
So the confidence interval is

\[
(0.5-0.4-1.64\sqrt{\frac{0.5 \cdot 0.5}{100} + \frac{0.4 \cdot 0.6}{100}}, 0.5-0.4+1.64\sqrt{\frac{0.5 \cdot 0.5}{100} + \frac{0.4 \cdot 0.6}{100}}) = (-0.0148, 0.2148).
\]

4.2.27. Let \(X_1, X_2, \ldots, X_n\) and \(Y_1, Y_2, \ldots, Y_m\) be two independent random samples from the respective normal distribution \(N(\mu_1, \sigma_1^2)\) and \(N(\mu_2, \sigma_2^2)\), where the four parameters are unknown. To construct a confidence interval for the ratio, \(\sigma_1^2/\sigma_2^2\), of the variances, form the quotient of the two independent \(\chi^2\) variables, each divided by its degrees of freedom, namely,

\[
F = \frac{(m-1)S_2^2}{\sigma_2^2}/(m-1) = \frac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2},
\]

where \(S_1^2\) and \(S_2^2\) are the respective sample variances.

(a) What kind of distribution does \(F\) have?

(b) From the appropriate table, \(a\) and \(b\) can be found so that \(P(F < b) = 0.975\) and \(P(a < F < b) = 0.95\).

(c) Rewrite the second probability statement as

\[
P\left[ a \frac{S_2^2}{\sigma_2^2} < b \frac{S_1^2}{\sigma_1^2} < b \frac{S_1^2}{\sigma_1^2} \right] = 0.95.
\]

The observed values, \(s_1^2\) and \(s_2^2\), can be inserted in these inequalities to provide a 95% confidence interval for \(\sigma_1^2/\sigma_2^2\).

Solution. (a) This is \(F\)-distribution with parameter \(m - 1\) and \(n - 1\).

(b) & (c) Note

\[
\frac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2} < b \iff \frac{\sigma_1^2}{\sigma_2^2} \cdot \frac{S_2^2}{S_1^2} < b \iff \frac{\sigma_1^2}{\sigma_2^2} < b \frac{S_1^2}{S_2^2},
\]

and

\[
\frac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2} > a \iff \frac{\sigma_1^2}{\sigma_2^2} \cdot \frac{S_2^2}{S_1^2} > a \iff \frac{\sigma_1^2}{\sigma_2^2} > a \frac{S_1^2}{S_2^2}.
\]

That’s why we can write \(P(a < \frac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2} < b) = 0.95\) as

\[
P\left( a \frac{S_2^2}{S_2^2} < \frac{\sigma_2^2}{\sigma_2^2} < b \frac{S_1^2}{S_2^2} \right) = 0.95.
\]