40.004 Statistics 2018: **Problem set 2 solutions**

1. (A twist on the German Tank Problem). Suppose that the enemy has tanks numbered $0, 1, 2, \ldots, N$. You observe n of the tanks *with* replacement at random and note down their numbers. Using the sample mean of these numbers, find an unbiased estimator for the total number of tanks (with justification for your claim).

   Solution: Note that $\mathbb{E}(\overline{X}) = \big(\mathbb{E}(X_1) + \mathbb{E}(X_2) + \cdots + \mathbb{E}(X_n)\big)/n$, which by symmetry equals $\mathbb{E}(X_1)$; here X_i denotes the number on the ith observed tank. Then

   $$\mathbb{E}(\overline{X}) = \mathbb{E}(X_1) = \frac{1}{N+1}\big(0 + 1 + 2 + \cdots + N\big) = \frac{N(N+1)/2}{N+1} = \frac{N}{2}.$$

   Note that the *total number* of tanks is actually $N+1$; this is the quantity we want to estimate. By linearity of expectation, $\mathbb{E}(2\overline{X}+1) = 2\dfrac{N}{2} + 1 = N + 1$. Thus $2\overline{X}+1$ is an unbiased estimator for the total number of tanks.

2. Let S^2 denote the sample variance computed from a random sample of size n from a $\mathcal{N}(\mu, \sigma^2)$ distribution. Find the probability that the sample variance S^2 exceeds the true variance σ^2 by a factor of two, i.e., $\Pr(S^2 > 2\sigma^2)$ when $n = 8, 17, 21$. Comment on your results. You may use R or Excel or a standard table in the book to find the probabilities.

   Solution: We know that if the random sample n is from $\mathcal{N}(\mu, \sigma^2)$ then $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$. Hence we have

   $$\Pr(S^2 > 2\sigma^2) = \Pr\left(\frac{(n-1)S^2}{\sigma^2} > 2(n-1)\right) = \Pr(W_{n-1} > 2(n-1)).$$

   where $W_{n-1} \sim \chi^2_{(n-1)}$. Using R we have

   $$\begin{aligned} n = 8: &\quad \Pr(S^2 > 2\sigma^2) = 0.05, \\ n = 17: &\quad \Pr(S^2 > 2\sigma^2) = 0.01, \\ n = 21: &\quad \Pr(S^2 > 2\sigma^2) = 0.005. \end{aligned}$$

   The probability that S^2 exceeds the true variance by a factor of two decreases as you increase the sample size. This is because our estimate of σ^2 improves as we have larger and larger samples.

3. A random sample of size 100, drawn from a normal distribution, has sample mean $\overline{x} = 16.3$.

   (a) Calculate the 95% two-sided confidence interval for μ, if $\sigma = 6$.

   (b) Calculate the 95% two-sided confidence interval for μ, if $s = 6$ and σ is unknown.

   (c) Calculate the upper and lower 95% one-sided confidence intervals for μ, if $s = 6$ and σ is unknown.

   (d) Why is the confidence interval in (b) wider than the CI in (a)?

   Solution:

   (a) We have $\bar{x} = 16.3$, $\alpha = 0.05$ and $n = 100$; the CI is given by

   $$\left[\bar{x} - z_{1-\alpha/2}\,\frac{\sigma}{\sqrt{n}},\ \bar{x} + z_{1-\alpha/2}\,\frac{\sigma}{\sqrt{n}}\right] = [15.124, 17.476].$$

(b) The CI is given by

$$\left[\bar{x} - t_{n-1,\,1-\alpha/2}\,\frac{s}{\sqrt{n}},\ \bar{x} + t_{n-1,\,1-\alpha/2}\,\frac{s}{\sqrt{n}}\right] = [15.109, 17.491].$$

(c) Upper 90% CI:

$$\left(-\infty,\ \bar{x} + t_{n-1,\,1-\alpha}\,\frac{s}{\sqrt{n}}\right] = (-\infty,\ 17.296].$$

Lower 90% CI:

$$\left[\bar{x} - t_{n-1,\,1-\alpha}\,\frac{s}{\sqrt{n}},\ \infty\right) = [15.303,\ \infty).$$

(d) The wider confidence interval reflects the greater uncertainty in (b) since the true value of σ is unknown and is estimated using s. Mathematically, this is a consequence of the fact that a t-distribution has heavier tails than the standard normal distribution.

4. Let $X_1, \ldots, X_n$ be a random sample from a distribution with mean μ and variance σ^2. Show that $\mathbb{E}(\overline{X}^2)$ is a biased estimator of μ^2.

   Solution: Note that

$$\frac{\sigma^2}{n} = \mathsf{Var}(\overline{X}) = \mathbb{E}(\overline{X}^2) - [\mathbb{E}(\overline{X})]^2.$$

   Hence

$$\mathbb{E}(\overline{X}^2) = [\mathbb{E}(\overline{X})]^2 + \mathsf{Var}(\overline{X}) = \mu^2 + \frac{\sigma^2}{n}.$$

   Unless $\sigma^2 = 0$, which would imply that X is a degenerate random variable taking the value μ almost everywhere, we have $\overline{X}^2$ to be a biased estimator of μ^2.

5. In each of the following cases, state the two competing hypotheses that should be tested and specify which would you set up as the null hypothesis and which one as the alternative hypothesis. Explain your choice briefly.

   (a) A consumer watchdog group suspects that a yogurt advertised to be 98% fat free has actually a higher fat content. The group plans to measure the fat contents of 25 yogurt cups (each containing 170 grams) to verify its suspicion.

   (b) It is claimed that cloud seeding is an effective technique to increase precipitation.

   Solution:

   (a) $\mathbf{H}_0 : \mu = 3.4$ vs $\mathbf{H}_\mathsf{A} : \mu > 3.4$ where μ is the fat content per yogurt cup. The choice of $\mathbf{H}_0$ is the status quo which is what is claimed by the yogurt company and the alternative hypothesis is purported by the consumer watchdog group.

   (b) The two hypotheses are: $\mathbf{H}_0$: cloud seeding is ineffective (status quo, no change in precipitation) vs. $\mathbf{H}_\mathsf{A}$: cloud seeding is effective (precipitation increases).

6. Consider testing $\mathbf{H}_0 : \mu = 0$ vs $\mathbf{H}_\mathsf{A} : \mu \neq 0$ based on a random sample of size n from a $\mathcal{N}(\mu, 1)$ distribution.

   (a) Calculate the p-values for the following three cases:
   (i) $\overline{x} = 0.1$, $n = 100$; (ii) $\overline{x} = 0.1$, $n = 400$; (iii) $\overline{x} = 0.1$, $n = 900$.

   (b) Given the significance level $\alpha = 0.01$, conduct the hypothesis tests for the three cases in (a).

   Solution: (a) To compute the p-value, we first assume that $\mathbf{H}_0$ is true, that is, $\mu = 0$.

   Then for (i), the p-value is

$$\Pr\left(|Z| \geq \left|\frac{0.1 - 0}{1/\sqrt{100}}\right|\right) = \Pr(|Z| \geq 1) = 0.3173;$$

   for (ii), $\Pr(|Z| \geq 2) = 0.0455$; for (iii), $\Pr(|Z| \geq 3) = 0.0027$.

   Note that the p-values are two-sided, since the alternative hypothesis is two-sided.

   (b) For (i) and (ii), the p-value is greater than α, so we do not reject $\mathbf{H}_0$; for (iii), the p-value is smaller than α, so we reject $\mathbf{H}_0$.

7. A tire company has developed a new tread design. To determine the newly designed tire has a mean of 60,000 miles or more, a random sample of 16 prototype tires are tested. The mean life for this sample is 60,758 miles. Assume that the tire life is normally distributed with unknown μ and standard deviation $\sigma = 1500$ miles. Test the hypothesis $\mathbf{H}_0 : \mu = 60,000$ vs. $\mathbf{H}_\mathsf{A} : \mu > 60,000$.

   (a) Compute the test statistic and the p-value. Based on the p-value, state whether $\mathbf{H}_0$ can be rejected at $\alpha = 0.01$.

   (b) What is the power of the 0.01-level test in (a) if the true mean life for the new tread design is $61,000$ miles?

   (c) Suppose that at least 90% power is needed to identify a tread design that has the mean life of $61,000$ miles. How many tires should be tested?

   Solution:

   (a) Let $x_1, \ldots, x_{16}$ be the observed values, $\overline{x}$ the sample mean and s^2 the sample variance. Then the test statistics is given by:

$$z = \frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{60758 - 60000}{1500/\sqrt{16}} = 2.021.$$

   The one-sided p-value is given by

$$\Pr(Z > z) = 1 - \Phi(2.021) = 0.022.$$

   Since $0.022 > 0.01$, we cannot reject $\mathbf{H}_0$ at level $\alpha = 0.01$. This means there is not enough evidence that the mean tire life for this sample differs from the average tire life for the old tread design.

   (b) If $\mu = 61000$, the power for this one-sided test is given by (formula was proved in class)

$$\begin{aligned}\pi(61000) &= \Pr(\text{reject } \mathbf{H}_0 | \mu = 61000) \\ &= \Phi\left(-z_\alpha + \frac{(\mu - \mu_0)\sqrt{n}}{\sigma}\right) \\ &= \Phi\left(-2.326 + \frac{(61000 - 60000)\sqrt{16}}{1500}\right) \\ &= \Phi(0.3407) = 0.633.\end{aligned}$$

(c) To assure 90% power in detecting a mean wear of 61000 miles, we use $\beta = 1 - \text{Power} = 0.1$. Then, again using formula from class we have

$$
\begin{aligned}
n &\approx \left[\frac{(z_\alpha + z_\beta)\sigma}{\mu - \mu_0}\right]^2 \\
&\approx \left[\frac{(2.326 + 1.282) \times 1500}{61000 - 60000}\right]^2 = 29.29.
\end{aligned}
$$

Hence, 30 tires should be tested.

8. Two methods of measuring the atomic weight of carbon (the nominal atomic weight is 12) yielded the following results.

Method 1	12.0129	12.0072	12.0064	12.0054	12.0016
	11.9853	11.9949	11.9985	12.0077	12.0061
Method 2	12.0318	12.0246	12.0069	12.0006	12.0075

(a) Test $\mathbf{H}_0 : \mu_1 = \mu_2$ vs. $\mathbf{H}_\mathsf{A} : \mu_1 \neq \mu_2$ at $\alpha = 0.05$, assuming $\sigma_1^2 = \sigma_2^2$. What is your conclusion?

(b) Repeat (a) without assuming $\sigma_1^2 = \sigma_2^2$. Compare the results.

Solution: The two data sets are $x_1, \ldots, x_{n_1}$ and $y_1, \ldots, y_{n_2}$. Hence from the data, we have $n_1 = 10, n_2 = 5$,

$$
\overline{x} = 12.0026, s_1 = 0.0079, \overline{y} = 12.0143, s_2 = 0.0132.
$$

(a) Assume $\sigma_1^2 = \sigma_2^2$. Using the above we see that the pooled standard deviation is

$$
s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} = 0.0099.
$$

Hence the test-statistic is

$$
t = \frac{\overline{x} - \overline{y}}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = -2.162.
$$

Since $|t| > t_{10+15-2,\alpha/2} = t_{23,0.025} = 2.16$, we reject $\mathbf{H}_0$ and conclude that there are significant differences between the methods (albeit the fact that we do this very marginally here).

(b) Assume $\sigma_1^2 \neq \sigma_2^2$. We compute the degrees of freedom of the required t-distribution here. Note that

$$
\omega_1 = \frac{s_1^2}{n_1} = 0.00000624, \omega_2 = \frac{s_2^2}{n_2} = 0.00001742.
$$

Hence the degrees of freedom here is given by

$$
\nu = \frac{(\omega_1 + \omega_2)^2}{\frac{\omega_1^2}{n_1 - 1} + \frac{\omega_2^2}{n_2 - 1}} = 5.481.
$$

The test-statistic is

$$
t = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = -1.817.
$$

Since $|t| < t_{0.025,5.48} = 2.504$, we cannot reject $\mathbf{H}_0$ at level $\alpha = 0.05$. Interestingly, this is the opposite conclusion from what we got for the case where we assumed equal variances.

9. A restaurant purchased a new oven, which is hoped to have more even heating than the old oven. By testing 9 locations inside each oven on the same temperature setting, it is found that the sample standard deviation for the temperature in the old oven is $s_1 = 2.3$, while that for the new oven is $s_2 = 1.1$. Set up a hypothesis test with $\alpha = 0.05$ to check whether the new oven indeed provides more even heating.

   Solution: Note that we are testing whether *two* populations have the same variance (so the F distribution is required); this problem cannot be reduced to testing the variance of a single population, since we do not know the true variance for either population.

   We have $\mathbf{H}_\mathsf{A} : \sigma_1 = \sigma_2$, $\mathbf{H}_\mathsf{A} : \sigma_1 > \sigma_2$, or equivalently, $\mathbf{H}_0 : \sigma_1^2/\sigma_2^2 = 1$, $\mathbf{H}_\mathsf{A} : \sigma_1^2/\sigma_2^2 > 1$.

   Assuming normality of the samples. Let S_1 and S_2 denote the random variables whose observed values are s_1 and s_2. Then we know that $\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$ is an $F_{8,\,8}$ random variable. Under $\mathbf{H}_0 = \sigma_1 = \sigma_2$, we should have

$$0.95 = \Pr\left(\frac{S_1^2}{S_2^2} \le f_{8,\,8,\,0.95}\right) = \Pr\left(\frac{S_1^2}{S_2^2} \le 3.438\right).$$

   We observe $\dfrac{s_1^2}{s_2^2} = 4.372$. Rearranging, and using $s_1 = 2.3$, $s_2 = 1.1$. Hence we reject $\mathbf{H}_0$. Alternatively we find that the lower 95% CI for σ_1^2/σ_2^2 is $[1.27,\ \infty)$ which does not contain 1. Thus we may conclude that the new oven indeed provides more even heating.

10. A person claims to be able to taste whether tea or milk was added first to a cup of English tea. To test her claim, 12 cups of visually indistinguishable tea are prepared, of which 6 of the cups are prepared tea-first, the other 6 milk-first. Being aware of this experimental setup, she would always try to pick 6 of the cups as tea-first, and the other 6 as milk-first. After tasting each cup of tea, she correctly identifies 5 of the tea-first cups (making 1 mistake), and 5 of the milk-first cups (also making 1 mistake).

    Compute the p-value, that is, the probability that one can do at least as well as her by guessing, and hence perform a hypothesis test at the $\alpha = 0.05$ level.

    Solution: Since the *tea-tasting lady* always picks exactly 6 of the cups as tea-first, and the rest as milk-first, if she identifies k of the tea-first cups correctly, then she will automatically get k of the milk-first cups right.

    So, to do at least as well as her means to either (i) correctly identify 5 of the tea-first cups (and hence also 5 of the milk-first cups); or (ii) correctly identify 6 of the tea-first cups (and hence also 6 of the milk-first cups) – that is, identify everything correctly.

    There are $\binom{12}{6} = 924$ ways to pick 6 cups from the 12 cups (ignoring the order in which they are chosen). The probability of getting (ii) by random guessing is just $1/924$. The probability of (i) is $\binom{6}{5}\binom{6}{1}/924$, as the $\binom{6}{5}$ accounts for the number of ways to get 5 of the 6 tea-first cups right, and the $\binom{6}{1}$ accounts for the number of ways to get 1 of the tea-first cups *wrong*.

    Adding up the two probabilities, we find that the p-value is $37/924 \approx 0.040 < \alpha$, therefore we conclude that there is significant evidence in support of her claim.