MA20033 - Solution Sheet Eight

Simon Shaw
s.c.shaw@maths.bath.ac.uk

2004/05 Semester I

1. **Suppose that we are going to gather a sample of twenty independent observations from a $N(\mu, \sigma^2 = 1.0)$ distribution in order to make inferences about μ. Suppose we are interested in the null hypothesis $H_0 : \mu = 10$.**

   (a) **State the two-tailed alternative hypothesis and find the appropriate critical values for a test of significance 0.05.**

   The alternative hypothesis is $H_1 : \mu \neq 10$. The critical region is $C^* = \{(x_1, \ldots, x_n) : \overline{x} \leq k_2,\ \overline{x} \geq k_1\}$ where the critical values k_1 and k_2, for a test of significance 0.05, are given by

   $$k_1 \ = \ 10 + 1.96\frac{1}{\sqrt{20}} \ = \ 10.4383, \ k_2 \ = \ 10 - 1.96\frac{1}{\sqrt{20}} \ = \ 9.5617$$

   as when H_0 is true, $\overline{X} \sim N(10, 1/20)$ and $z_{0.975} = 1.96$.

   i. **What is the p-value of this test when we observe a sample mean $\bar{x} = 10.41$?**

   The p-value of this test when we observe a sample mean $\overline{x} = 10.41$ is the probability, under H_0, of getting a sample mean at least as extreme as 10.41, that is getting a sample mean at least as large as 10.41 or smaller than or equal to $10 - 0.41 = 9.59$. We note that if $\overline{x} = 10.41$ then there is insufficient evidence to reject H_0 at the 5% significance level so we note that the p-value > 0.05.

   $$\begin{aligned} \text{p-value} &= P\{\overline{X} \leq 9.59 \cup \overline{X} \geq 10.41 \mid \overline{X} \sim N(10, 1/20)\} \\ &= 2P\{\overline{X} \geq 10.41 \mid \overline{X} \sim N(10, 1/20)\} \text{ by symmetry} \\ &= 2P\{Z \geq \sqrt{20}(10.41 - 10)\} \\ &= 2\{1 - \Phi(1.83)\} \ = \ 2(1 - 0.9664) \ = \ 0.0672. \end{aligned}$$

   Thus, we reject H_0 for all tests with significance greater than or equal to 0.0672.

   ii. **Evaluate the power at $\mu = 10.5$.**

   $$\begin{aligned} \pi(10.5) \ &= \ P\{\overline{X} \geq 10 + 1.96/\sqrt{20} \mid \overline{X} \sim N(10.5, 1/20)\} + \\ & \qquad P\{\overline{X} \leq 10 - 1.96/\sqrt{20} \mid \overline{X} \sim N(10.5, 1/20)\} \\ &= \ P\{Z \geq \sqrt{20}(10 - 10.5) + 1.96\} + P\{Z \leq \sqrt{20}(10 - 10.5) - 1.96\} \\ &= \ 1 - \Phi(-0.28) + \Phi(-4.20) \\ &= \ \Phi(0.28) \ = \ 0.6103. \end{aligned}$$

(b) **State the upper one-tailed alternative hypothesis and find the appropriate critical values for a test of significance 0.05.**

The alternative hypothesis is $H_1 : \mu > 10$. The critical region is $C^* = \{(x_1, \ldots, x_n) : \overline{x} \geq k_1\}$ where the critical value k_1, for a test of significance 0.05, is given by

$$k_1 \;=\; 10 + 1.645\frac{1}{\sqrt{20}} \;=\; 10.3678$$

as when H_0 is true, $\overline{X} \sim N(10, 1/20)$ and $z_{0.95} = 1.645$.

i. **What is the p-value of this test when we observe a sample mean $\overline{x} = 10.41$?**

The p-value of this test when we observe a sample mean $\overline{x} = 10.41$ is the probability, under H_0, of getting a sample mean as least as extreme as 10.41, that is getting a sample mean greater than or equal to 10.41. Notice that $10.41 > 10.3678$ so we reject H_0 at the 5% level. Hence, the p-value < 0.05.

$$\begin{aligned} \text{p-value} &= P\{\overline{X} \geq 10.41 \mid \overline{X} \sim N(10, 1/20)\} \\ &= P\{Z \geq \sqrt{20}(10.41 - 10)\} \\ &= 1 - \Phi(1.83) \;=\; 1 - 0.9664 \;=\; 0.0336. \end{aligned}$$

Thus, we reject H_0 for all tests with significance greater than or equal to 0.0336.

ii. **Evaluate the power at $\mu = 10.5$.**

$$\begin{aligned} \pi(10.5) &= P\{\overline{X} \geq 10 + 1.645/\sqrt{20} \mid \overline{X} \sim N(10.5, 1/20)\} \\ &= P\{Z \geq \sqrt{20}(10 - 10.5) + 1.645\} \\ &= 1 - \Phi(-0.59) \;=\; \Phi(0.59) \;=\; 0.7224. \end{aligned}$$

2. **The Food and Nutrition Board of the National Academy of Sciences (USA) suggests that the RDA of iron for adult females under the age of 51 is 18mg. The following iron intakes, in mg, during a 24-hour period were observed for 45 randomly selected adult females in this age group:**

15.0	18.1	14.4	14.6	10.9	18.1	18.2	18.3	15.0
16.0	12.6	16.6	20.7	19.8	11.6	12.8	15.6	11.0
15.3	9.4	19.5	18.3	14.5	16.6	11.5	16.4	12.5
14.6	11.9	12.5	18.6	13.1	12.1	10.7	17.3	12.4
17.0	6.3	16.8	12.5	16.3	14.7	12.7	16.3	11.5

Under the assumption that each observation represents the realisation of a $N(\mu, \sigma^2)$ random quantity, derive a random interval which contains σ^2 with probability 0.95. Evaluate a 95% confidence interval for σ^2. Use may use the fact that $\sum_{i=1}^{45} x_i = 660.6$, $\sum_{i=1}^{45} x_i^2 = 10115.88$, $\chi^2_{44,0.975} = 27.574$ and $\chi^2_{44,0.025} = 64.201$.

Under the assumption that each of the 45 observations represent a realisation of a

$N(\mu, \sigma^2)$ random quantity, then S^2 is an unbiased estimator of σ^2 with $(45-1)S^2/\sigma^2 \sim \chi^2_{44}$. Therefore,

$$P(\chi^2_{44,0.975} < \chi^2_{44} < \chi^2_{44,0.025}) \;=\; 0.95 \;\Rightarrow\; P(27.574 < \chi^2_{44} < 64.201) \;=\; 0.95.$$

Using the pivot $44S^2/\sigma^2 \sim \chi^2_{44}$ we have that

$$\begin{aligned} P(27.574 < 44S^2/\sigma^2 < 64.201 \mid \sigma^2) &= 0.95 \\ \Rightarrow P(44S^2/64.201 < \sigma^2 < 44S^2/27.574 \mid \sigma^2) &= 0.95. \end{aligned}$$

Thus, $(44S^2/64.201, 44S^2/27.574)$ is a random interval which contains σ^2 with probability 0.95. A realisation of this, $(44s^2/64.201, 44s^2/27.574)$, is a 95% confidence interval for σ^2. For the RDA data we have

$$\overline{x} \;=\; \frac{1}{n}\sum_{i=1}^{n} x_i \;=\; \frac{660.6}{45} \;=\; 14.68,$$

and

$$\begin{aligned} s^2 \;=\; \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2\right) &= \frac{1}{44}\left\{10115.88 - 45\left(\frac{660.6}{45}\right)^2\right\} \\ &= \frac{418.272}{44} \;=\; 9.5062. \end{aligned}$$

so that our 95% confidence interval for σ^2 is

$$(44s^2/64.201, 44s^2/27.574) = (418.272/64.201, 418.272/27.574) = (6.5150, 15.1691).$$

3. **Under the same assumptions as question 2, and using the same data, test the hypotheses**

$$H_0 : \sigma^2 = 10 \quad \text{versus} \quad H_1 : \sigma^2 \neq 10$$

Express the p-value of this test as a probability statement involving the test statistic (the tables which you have will not give the necessary quantiles to evaluate this p-value, but you could investigate using Minitab or R to find it).

We use S^2 as the test statistic and the critical region is $C^* = \{(x_1, \ldots, x_n) : s^2 \leq k_2,\ s^2 \geq k_1\}$ where the critical values k_1 and k_2 are chosen to ensure a test of significance α. Under H_0, $44S^2/10 \sim \chi^2_{44}$ and so we choose the critical values so that

$$P(S^2 \leq k_2 \cup S^2 \geq k_1 \mid 44S^2/10 \sim \chi^2_{44}) \;=\; \alpha.$$

We allocate $\alpha/2$ to each tail error so that

$$\begin{aligned} P(S^2 \geq k_1 \mid 44S^2/10 \sim \chi^2_{44}) &= \frac{\alpha}{2} \\ \Rightarrow P(44S^2/10 \geq 44k_1/10 \mid 44S^2/10 \sim \chi^2_{44}) &= \frac{\alpha}{2}. \end{aligned}$$

Hence, $k_1 = 10\chi^2_{44,\alpha/2}/44$. For the lower tail,

$$\begin{aligned} P(S^2 \leq k_2 \mid 44S^2/10 \sim \chi^2_{44}) &= \frac{\alpha}{2} \\ \Rightarrow P(44S^2/10 \leq 44k_2/10 \mid 44S^2/10 \sim \chi^2_{44}) &= \frac{\alpha}{2}. \end{aligned}$$

Thus, $k_2 = 10\chi^2_{44,1-\alpha/2}/44$. For a test of significance 0.05 we have $k_1 = 10(64.201)/44 = 14.5911$ and $k_2 = 10(27.574)/44 = 6.2668$. We observe $s^2 = 9.5062$ so don't reject H_0 at the 5% significance level. Notice that we could deduce this immediately from the 95% confidence interval for σ^2, $(6.5150, 15.1691)$ obtained in question 2., as 9.5062 is contained in this interval.

The p-value of this test is a probability that we observe a test statistic as least as extreme than that observed when H_0 is true. Notice that $E(S^2|\sigma^2 = 10) = 10$ as S^2 is an unbiased estimator of σ^2 which equals 10 under H_0. Our observed value $s^2 = 418.272/44 < 10$ and, given that this is a two-tailed test,

$$\begin{aligned} \text{p-value} &= 2P(S^2 \leq 418.272/44 | 44S^2/10 \sim \chi^2_{44}) \\ &= 2P(\chi^2_{44} \leq 418.272/10) \\ &= 2P(\chi^2_{44} \leq 41.8272) \\ &= 2(0.4348) \ = \ 0.8696, \end{aligned}$$

where, using R, $P(\chi^2_{44} \leq 41.8272) = pchisq(41.8272, 44) = 0.4348$. Thus, we reject H_0 at significance α for $\alpha \geq 0.8696$.