# MA20033 - Solution Sheet Seven

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

2004/05 Semester I

1. **Suppose $X_1, \ldots, X_{10}$ are independent and identically distributed $N(\mu, \sigma^2 = 25)$ random quantities.**

   (a) **Carry out the following one-sided test by stating a test statistic and finding the critical value of the test:**

   $$H_0 : \mu = 105 \quad \text{versus} \quad H_1 : \mu > 105$$

   A test statistic is $\overline{x}$ and the critical region of the uniformly most powerful test is $C^* = \{(x_1, \ldots, x_n) : \overline{x} \geq k\}$. For a test of significance $\alpha$, the critical value

   $$k = 105 + z_{(1-\alpha)} \frac{5}{\sqrt{10}}$$

   as when $H_0$ is true, $\overline{X} \sim N(105, 25/10)$.

   (b) **Carry out this test at the 5% significance level when a sample mean $\overline{x} = 108$ is obtained.**

   For a test at the 5% significance level, $z_{0.95} = 1.645$ and $k = 105 + 1.645\sqrt{2.5} = 107.60$. Thus, for $\overline{x} = 108$, we reject $H_0$ in favour of $H_1$ at the 5% significance level. There is sufficient evidence to suggest that the mean is greater than 105.

   (c) **Evaluate the power of this test at the points $\mu = 106$ and $\mu = 110$.**

   The power is the probability of rejecting $H_0$ when $H_1$ is true. In this case, the power is a function of $\mu$, $\pi(\mu)$, and given by

   $$\begin{aligned} \pi(\mu) &= P\{\overline{X} \geq 105 + 1.645\sqrt{2.5} \mid \overline{X} \sim N(\mu, 2.5)\} \\ &= P\{Z \geq (105 - \mu)/\sqrt{2.5} + 1.645 \mid Z \sim N(0,1)\} \\ &= 1 - \Phi\left(\frac{105 - \mu}{\sqrt{2.5}} + 1.645\right). \end{aligned}$$

   Thus, for $\mu = 106$,

   $$\begin{aligned} \pi(106) &= 1 - \Phi\left(\frac{105 - 106}{\sqrt{2.5}} + 1.645\right) \\ &= 1 - \Phi(1.01) = 1 - 0.8438 = 0.1562. \end{aligned}$$

For $\mu = 110$,

$$\begin{aligned}\pi(110) &= 1 - \Phi\left(\frac{105-110}{\sqrt{2.5}} + 1.645\right) \\ &= 1 - \Phi(-1.52) = \Phi(1.52) = 0.9357.\end{aligned}$$

Note that we expect $\pi(110) > \pi(106) > 0.05$. The power function is a monotonically increasing curve which tends to 1 as $\mu$ tends to infinity (and tends to zero as $\mu$ tends to minus infinity, although really we are only interested in $\mu > 105$ given this $H_1$). It takes the value 0.05 at $\mu = 105$, i.e. it equals the significance level of the test at the $H_0$ value of $\mu$.

2. **Again suppose $X_1, \ldots, X_{10}$ are independent and identically distributed $N(\mu, \sigma^2 = 25)$ random quantities.**

   (a) **Carry out the following two-sided test by stating a test statistic and finding the critical value of the test:**

   $$H_0 : \mu = 105 \quad \text{versus} \quad H_1 : \mu \neq 105$$

   A test statistic is $\overline{x}$. To obtain a critical region of a test with significance $\alpha$, we combine the critical regions of the uniformly most powerful test of $H_1 : \mu > 105$ and $H_1 : \mu < 105$, both at significance level $\alpha/2$. Thus, $C^* = \{(x_1, \ldots, x_n) : \overline{x} \leq k_2, \overline{x} \geq k_1\}$ where the critical values $k_1$ and $k_2$ are given by

   $$k_1 = 105 + z_{(1-\frac{\alpha}{2})}\frac{5}{\sqrt{10}}, \quad k_2 = 105 - z_{(1-\frac{\alpha}{2})}\frac{5}{\sqrt{10}}$$

   as when $H_0$ is true, $\overline{X} \sim N(105, 25/10)$.

   (b) **Carry out this test at the 5% significance level when a sample mean $\overline{x} = 108$ is obtained.**

   When $\alpha = 0.05$, $z_{0.975} = 1.96$. Hence, $k_1 = 105 + 1.96\sqrt{2.5} = 108.10$ and $k_2 = 105 - 1.96\sqrt{2.5} = 101.90$. If $\overline{x} = 108$ then $101.90 < 108 < 108.10$. We do not reject $H_0$. There is not enough evidence at the 5% level to conclude that the mean is not 105.

   (c) **Evaluate the power of this test at the points $\mu = 106$ and $\mu = 110$.**

   The power is the probability of rejecting $H_0$ when $H_1$ is true. In this case, the power is a function of $\mu$, $\pi(\mu)$, and given by

   $$\begin{aligned}\pi(\mu) &= P\{\overline{X} \geq 105 + 1.96\sqrt{2.5}, \overline{X} \leq 105 - 1.96\sqrt{2.5} \mid \overline{X} \sim N(\mu, 2.5)\} \\ &= P\{Z \geq (105-\mu)/\sqrt{2.5} + 1.96 \mid Z \sim N(0,1)\} + \\ &\qquad P\{Z \leq (105-\mu)/\sqrt{2.5} - 1.96 \mid Z \sim N(0,1)\} \\ &= 1 - \Phi\left(\frac{105-\mu}{\sqrt{2.5}} + 1.96\right) + \Phi\left(\frac{105-\mu}{\sqrt{2.5}} - 1.96\right).\end{aligned}$$

   Thus, for $\mu = 106$,

   $$\begin{aligned}\pi(106) &= 1 - \Phi\left(\frac{105-106}{\sqrt{2.5}} + 1.96\right) + \Phi\left(\frac{105-106}{\sqrt{2.5}} - 1.96\right) \\ &= 1 - \Phi(1.33) + \Phi(-2.59) \\ &= 1 - 0.9082 + (1 - 0.9952) = 0.0966.\end{aligned}$$

For $\mu = 110$,

$$\begin{aligned}\pi(110) &= 1 - \Phi\left(\frac{105-110}{\sqrt{2.5}} + 1.96\right) + \Phi\left(\frac{105-110}{\sqrt{2.5}} - 1.96\right) \\ &= 1 - \Phi(-1.20) + \Phi(-5.12) \\ &= 1 - (1 - 0.8849) + (1 - 1) \ = \ 0.8849.\end{aligned}$$

Note that we expect $\pi(110) > \pi(106) > 0.05$. The power is a symmetric curve, with the point of symmetry at $\mu = 105$ where the power equals the significance level of the test, i.e. 0.05. The power tends to 1 as $|\mu - 105|$ increases.

(d) **Would you have expected the power to be greater or smaller than that of the one-sided test?**

We would have expected the power to be smaller than that of the one-sided test. The reasoning is as follows. Notice that we could use the critical region $C^* = \{(x_1, \ldots, x_n) : \overline{x} \leq 105 - z_{(1-\frac{\alpha}{2})}\sqrt{2.5}, \overline{x} \geq 105 + z_{(1-\frac{\alpha}{2})}\sqrt{2.5}\}$ as a critical region for a test of $H_1 : \mu > 105$. However, for this test, in question 1.,we used the critical region corresponding to the uniformally most powerful test (derived from the Neyman-Pearson lemma). Thus, the test with critical region $C^* = \{(x_1, \ldots, x_n) : \overline{x} \geq 105 + z_{(1-\alpha)}\sqrt{2.5}\}$ has the largest power amongst all tests of significance $\alpha$ of $H_1 : \mu > 105$.

3. **Suppose $X_1, \ldots, X_n$ are independent and identically distributed $N(\mu, 1)$ random quantities, and you want a test with significance level 5% of the one-sided hypotheses**

$$H_0 : \mu = 0 \quad \text{versus} \quad H_1 : \mu < 0$$

(a) **Find the critical value $k$ such that the null hypothesis will be rejected if $\overline{x} < k$ (this critical value will depend on the sample size through the variance of $\overline{X}$).**

A test statistic is $\overline{x}$ and the critical region of the uniformly most powerful test is $C^* = \{(x_1, \ldots, x_n) : \overline{x} \leq k\}$. For a test of significance $\alpha$, the critical value

$$k = -1.645\frac{1}{\sqrt{n}}$$

as when $H_0$ is true, $\overline{X} \sim N(0, 1/n)$ and $z_{0.95} = 1.645$.

(b) **Suppose that you require your test to have power of at least 0.95 when $\mu$ is actually equal to -0.25. What is the probability that $H_0$ is not rejected when $\mu = -0.25$ (as a function of $n$)? So how big must the sample size $n$ be to meet this power requirement?**

The power is the probability of rejecting $H_0$ when $H_1$ is true. If $\mu = -0.25$, the power, $\pi(-0.25)$, is given by

$$\begin{aligned}\pi(-0.25) &= P\{\overline{X} \leq -1.645/\sqrt{n} \mid \overline{X} \sim N(-0.25, 1/n)\} \\ &= P\{Z \leq 0.25/\sqrt{1/n} - 1.645 \mid Z \sim N(0,1)\} \\ &= \Phi(0.25\sqrt{n} - 1.645).\end{aligned}$$

In order for the test to have power of at least 0.95 when $\mu = -0.25$, we require that

$$\begin{aligned} \Phi(0.25\sqrt{n} - 1.645) &\geq 0.95 \Rightarrow \\ 0.25\sqrt{n} - 1.645 &\geq 1.645 \Rightarrow \\ \sqrt{n} &\geq 8(1.645) = 13.16. \end{aligned}$$

Bearing in mind that $n$ is the sample size, and so takes integer values, we need $n \geq 174$ to achieve the desired power level.

4. **Suppose $X_1, \ldots, X_n$ are independent and identically distributed $N(\mu, \sigma^2)$ random quantities.**

   (a) **What is the sampling distribution of $S^2$?**

   We use the result that $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$, the $\chi^2$-distribution with $n-1$ degrees of freedom, to deduce the distribution of $S^2$.

   (b) **Set up a hypothesis test to determine whether the variability of some process has changed: Test the hypothesis $H_0 : \sigma^2 = 10$ against $H_1 : \sigma^2 \neq 10$, when a sample of size 20 yields an observed value of $s^2 = 13.8$.**

   Our critical region is $C^* = \{(x_1, \ldots, x_n) : s^2 < k_2, s^2 > k_1\}$ where $k_1$, $k_2$ are chosen so that

$$P(S^2 < k_2, S^2 > k_1 \mid H_0 \text{ true}) = \alpha$$

   to yield a test with significance $\alpha$. Our approach is to take

$$\begin{aligned} &P(S^2 < k_2 \mid H_0 \text{ true}) = \\ &\quad P\left\{ \frac{(20-1)S^2}{10} < \frac{(20-1)k_2}{10} \;\middle|\; \frac{(20-1)S^2}{10} \sim \chi^2_{20-1} \right\} = \frac{\alpha}{2} \end{aligned}$$

   so that $k_2 = \frac{10}{19}\chi^2_{19,1-\frac{\alpha}{2}}$ and

$$\begin{aligned} &P(S^2 > k_1 \mid H_0 \text{ true}) = \\ &\quad P\left\{ \frac{(20-1)S^2}{10} > \frac{(20-1)k_1}{10} \;\middle|\; \frac{(20-1)S^2}{10} \sim \chi^2_{20-1} \right\} = \frac{\alpha}{2} \end{aligned}$$

   so that $k_1 = \frac{10}{19}\chi^2_{19,\frac{\alpha}{2}}$.

   When $\alpha = 0.05$, $\chi^2_{19,0.975} = 8.907$ and $\chi^2_{19,0.025} = 32.852$ so $k_2 = 4.688$ and $k_1 = 17.290$. For a test with a significance level of 5%, the critical region is $C^* = \{(x_1, \ldots, x_n) : s^2 < 4.688, s^2 > 17.290\}$. We observe a value of $s^2 = 13.8$, which does not fall within the critical region. We conclude that we have insufficient evidence to reject $H_0$: we have insufficient evidence to suggest that the variability of the process has changed.

5. **The Neyman-Pearson lemma states that of all tests with significance $\alpha$ of the simple hypotheses**

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta = \theta_1$$

**the test which uses the critical region**

$$C^* \;=\; \{(x_1,\ldots,x_n) : \lambda(x_1,\ldots,x_n;\theta_0,\theta_1) \le k\},$$

**where** $\lambda(x_1,\ldots,x_n;\theta_0,\theta_1) = f(x_1,\ldots,x_n|\theta=\theta_0)/f(x_1,\ldots,x_n|\theta=\theta_1)$ **and** $k$ **is a constant chosen to ensure the significance is** $\alpha$**, is the one with the largest power. Denote** $(X_1,\ldots,X_n)$ **by** $\underline{X}$ **and** $(x_1,\ldots,x_n)$ **by** $\underline{x}$**.**

(a) **Show that for any set** $A \subset C^*$**,**

$$P(\underline{X} \in A \mid \theta=\theta_0) \;\le\; kP(\underline{X} \in A \mid \theta=\theta_1).$$

If $\underline{x} \in A \subset C^*$ then $0 \le f(\underline{x} \mid \theta=\theta_0) \le kf(\underline{x} \mid \theta=\theta_1)$. Thus,

$$\begin{aligned} P(\underline{X} \in A \mid \theta=\theta_0) &= \int_A f(\underline{x} \mid \theta=\theta_0)\,d\underline{x} \\ &\le \int_A kf(\underline{x} \mid \theta=\theta_1)\,d\underline{x} \\ &= kP(\underline{X} \in A \mid \theta=\theta_1). \end{aligned}$$

(b) **Show that for any set** $A \subset \overline{C^*}$ **(the compliment of** $C^*$**)**

$$P(\underline{X} \in A \mid \theta=\theta_0) \;>\; kP(\underline{X} \in A \mid \theta=\theta_1).$$

For $\overline{C^*}$, $\lambda(x_1,\ldots,x_n;\theta_0,\theta_1) > k$ so that $f(\underline{x} \mid \theta=\theta_0) > kf(\underline{x} \mid \theta=\theta_1) \ge 0$. Hence, for $A \subset \overline{C^*}$ we have

$$\begin{aligned} P(\underline{X} \in A \mid \theta=\theta_0) &= \int_A f(\underline{x} \mid \theta=\theta_0)\,d\underline{x} \\ &> \int_A kf(\underline{x} \mid \theta=\theta_1)\,d\underline{x} \\ &= kP(\underline{X} \in A \mid \theta=\theta_1). \end{aligned}$$

(c) **Let** $C$ **be some other critical region with significance level** $\alpha$**. Show that**

$$P(\underline{X} \in C^*|\theta) - P(\underline{X} \in C|\theta) \;=\; P(\underline{X} \in \{C^* \cap \overline{C}\}|\theta) - P(\underline{X} \in \{C \cap \overline{C^*}\}|\theta),$$

**where** $\overline{C}$ **denotes the compliment of** $C$**.**

We may partition $C^*$ into mutually exclusive sets: that which contains $C$ and that which doesn't contain $C$. That is, $C^* = \{C^* \cap C\} \cup \{C^* \cap \overline{C}\}$. Thus,

$$P(\underline{X} \in C^* \mid \theta) \;=\; P(\underline{X} \in \{C^* \cap C\} \mid \theta) + P(\underline{X} \in \{C^* \cap \overline{C}\} \mid \theta). \qquad (1)$$

Similarly, we may write $C = \{C \cap C^*\} \cup \{C \cap \overline{C^*}\}$ so that

$$P(\underline{X} \in C \mid \theta) \;=\; P(\underline{X} \in \{C \cap C^*\} \mid \theta) + P(\underline{X} \in \{C \cap \overline{C^*}\} \mid \theta). \qquad (2)$$

Subtracting (2) from (1) gives the result.

(d) **Hence show that**

$$P(\underline{X} \in C^* \mid \theta=\theta_1) - P(\underline{X} \in C \mid \theta=\theta_1) \;\ge\; 0.$$

**You have now proved the Neyman-Pearson lemma.**

From (c) we have that

$$\begin{aligned}&P(\underline{X} \in C^* \mid \theta = \theta_1) - P(\underline{X} \in C \mid \theta = \theta_1) \ = \\ &\qquad P(\underline{X} \in \{C^* \cap \overline{C}\} \mid \theta = \theta_1) - P(\underline{X} \in \{C \cap \overline{C^*}\} \mid \theta = \theta_1).\end{aligned}$$

Now, $C^* \cap \overline{C} \subset C^*$, so from (a)

$$P(\underline{X} \in \{C^* \cap \overline{C}\} \mid \theta = \theta_1) \quad \geq \quad \frac{1}{k} P(\underline{X} \in \{C^* \cap \overline{C}\} \mid \theta = \theta_0).$$

Also, $C \cap \overline{C^*} \subset \overline{C^*}$, so from (b)

$$P(\underline{X} \in \{C \cap \overline{C^*}\} \mid \theta = \theta_1) \quad > \quad \frac{1}{k} P(\underline{X} \in \{C \cap \overline{C^*}\} \mid \theta = \theta_0).$$

Hence,

$$\begin{aligned}&P(\underline{X} \in C^* \mid \theta = \theta_1) - P(\underline{X} \in C \mid \theta = \theta_1) \\ &> \quad \frac{1}{k}\{P(\underline{X} \in \{C^* \cap \overline{C}\} \mid \theta = \theta_0) - P(\underline{X} \in \{C \cap \overline{C^*}\} \mid \theta = \theta_0)\} \\ &= \quad \frac{1}{k}\{P(\underline{X} \in C^* \mid \theta = \theta_0) - P(\underline{X} \in C \mid \theta = \theta_0)\}\end{aligned}$$

Now, $C^*$ is a critical region with significance $\alpha$, so $P(\underline{X} \in C^* \mid \theta = \theta_0) = \alpha$. As $C$ is also a critical region with significance $\alpha$, $P(\underline{X} \in C \mid \theta = \theta_0) = \alpha$. Thus,

$$P(\underline{X} \in C^* \mid \theta = \theta_1) - P(\underline{X} \in C \mid \theta = \theta_1) \quad > \quad 0.$$

That is the probability of rejecting a false $H_0$ using $C^*$ is greater than that using $C$, i.e. the critical region $C^*$ has the largest power amongst all critical regions with significance $\alpha$.