# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 1**

**Question:**

Historical records from a large colony of squirrels show that the weight of squirrels is normally distributed with a mean of 1012 g. Following a change in the diet of squirrels, a biologist is interested in whether or not the mean weight has changed.
A random sample of 14 squirrels is weighed and their weights $x$, in grams, recorded.
The results are summarised as follows:

$$\sum x = 13\,700, \quad \sum x^2 = 13\,448\,750 .$$

Stating your hypotheses clearly test, at the 5% level of significance, whether or not there has been a change in the mean weight of the squirrels. **[E]**

**Solution:**

$\mathrm{H_0}: \mu = 1012 \quad \mathrm{H_1}: \mu \neq 1012$

$$\bar{x} = \frac{13\,700}{14} (= 978.57\ldots)$$

$$S_x^2 = \frac{13\,448\,750 - 14\bar{x}^2}{13} (= 3255.49)$$

$$t_{13} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{978.6 - 1012}{\frac{57.06}{\sqrt{14}}} = -2.19\ldots$$

$t_{13}$ (5%) two-tail critical value $= -2.160$

Significant result – there is evidence of a change in mean weight of squirrels

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 2**

**Question:**

A random sample $X_1, X_2, \ldots, X_{10}$ is taken from a population with mean $\mu$ and variance $\sigma^2$.

**a** Determine the bias, if any, of each of the following estimators of $\mu$.

$$\theta_1 = \frac{X_3 + X_4 + X_5}{3},$$

$$\theta_2 = \frac{X_{10} - X_1}{3},$$

$$\theta_3 = \frac{3X_1 + 2X_2 + X_{10}}{6}$$

**b** Find the variance of each of these estimators.

**c** State, giving reasons, which of these three estimators for $\mu$ is

**i** the best estimator,

**ii** the worst estimator. **[E]**

**Solution:**

**a** $\mathrm{E}(\theta_1) = \frac{\mathrm{E}\left((X_3) + \mathrm{E}(X_4) + \mathrm{E}(X_5)\right)}{3} = \frac{3\mu}{3} = \mu$ Bias $= 0$

$\mathrm{E}(\theta_2) = \frac{\mathrm{E}(X_{10}) - \mathrm{E}(X_1)}{3} = \frac{1}{3}(\mu - \mu) = 0$ Bias $= -\mu$

$\mathrm{E}(\theta_3) = \frac{3\mathrm{E}(X_1) + 2\mathrm{E}(X_2) + \mathrm{E}(X_{10})}{6} = \frac{3\mu + 2\mu + \mu}{6} = \mu$ Bias $= 0$

**b** $\mathrm{Var}(\theta_1) = \frac{1}{9}\left(\sigma^2 + \sigma^2 + \sigma^2\right) = \frac{\sigma^2}{3}$

$\mathrm{Var}(\theta_2) = \frac{1}{9}\left(\sigma^2 + \sigma^2\right) = \frac{2\sigma^2}{9}$

$\mathrm{Var}(\theta_3) = \frac{1}{36}\left[9\sigma^2 + 4\sigma^2 + \sigma^2\right] = \frac{14\sigma^2}{36} = \frac{7\sigma^2}{18}$

**c** Don't use $\theta_2$ as it is biased.

$\mathrm{Var}(\theta_1) = \frac{\sigma^2}{3} = \frac{6\sigma^2}{18}$

$\mathrm{Var}(\theta_3) = \frac{7\sigma^2}{18}$

**i** So choose $\theta_1$ as it is unbiased and has the smallest variance, to be the best estimator

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 3**

**Question:**

A random sample of 10 mustard plants had the following heights, in millimetres, after 4 days growth.

5.0, 4.5, 4.8, 5.2, 4.3, 5.1, 5.2, 4.9, 5.1, 5.0

Those grown previously had a mean height of 5.1 mm after 4 days. Using a 2.5% significance level, test whether or not the mean height of these plants is less than that of those grown previously.

(You may assume that the height of mustard plants after 4days follows a normal distribution.) **[E]**

**Solution:**

$\mathrm{H}_0 : \mu = 5.1, \mathrm{H}_1 : \mu < 5.1$

$\nu = 9$

Critical Region $t < -2.262$

$\bar{x} = 4.91$

$$s^2 = \frac{241.89 - 10 \times (4.91)^2}{9} = 0.0899$$

$s = 0.300$

$$t = \frac{4.91 - 5.1}{\frac{0.3}{\sqrt{10}}} = -2.00$$

There is no evidence to suggest that the mean height is less than those grown previously

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 4**

**Question:**

A mechanic is required to change car tyres. An inspector timed a random sample of 20 tyre changes and calculated the unbiased estimate of the population variance to be 6.25 minutes$^2$. Test, at the 5% significance level, whether or not the standard deviation of the population of times taken by the mechanic is grater than 2 minutes. State your hypotheses clearly.

**[E]**

**Solution:**

$H_0 : \sigma^2 = 4; H_1 : \sigma^2 > 4$

$\nu = 19, X^2_{19}(0.05) = 30.144$

$\frac{(n-1)S^2}{\sigma^2} = \frac{19 \times 6.25}{4} = 29.6875$

Since $29.6875 < 30.144$ there is insufficient evidence to reject $H_0$.

There is insufficient evidence to suggest that the standard deviation is greater than 2.

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 5**

**Question:**

The value of orders, in £, made to a firm over the internet has distribution $\text{N}(\mu, \sigma^2)$.

A random sample of $n$ orders is taken and $\bar{X}$ denotes the sample mean.

**a** Write down the mean and variance of $\bar{X}$ in terms of $\mu$ and $\sigma^2$.

A second sample of $m$ orders is taken and $\bar{Y}$ denotes the mean of this sample.

An estimator of the population mean is given by $U = \dfrac{n\bar{X} + m\bar{Y}}{n+m}$.

**b** Show that $U$ is an unbiased estimator for $\mu$.

**c** Show that the variance of $U$ is $\dfrac{\sigma^2}{n+m}$.

**d** State which of $\bar{X}$ or $U$ is a better estimator for $\mu$. Give a reason for your answer.

**[E]**

**Solution:**

**a** $\text{E}(\bar{X}) = \mu$

$$\text{Var}(\bar{X}) = \text{Var}\left(\frac{X_1 + X_2 + X_3 + \ldots + X_n}{n}\right) = \frac{\sigma^2}{n}$$

**b**
$$\text{E}(U) = \frac{1}{n+m}\left(n\text{E}\left(\bar{X}\right) + m\text{E}\left(\bar{Y}\right)\right)$$
$$= \frac{1}{n+m}(n\mu + m\mu)$$
$$= \mu \Rightarrow U \text{ is unbiased}$$

**c**
$$\text{Var}\left(\bar{Y}\right) = \frac{\sigma^2}{m}$$
$$\text{Var}(U) = \frac{n^2\text{Var}\left(\bar{X}\right) + m^2\text{Var}\left(\bar{Y}\right)}{(n+m)^2}$$
$$= \frac{n^2\dfrac{\sigma^2}{n} + m^2\dfrac{\sigma^2}{m}}{(n+m)^2}$$
$$= \frac{n\sigma^2 + m\sigma^2}{(n+m)^2}$$
$$= \frac{\sigma^2}{n+m}$$

**d** $\dfrac{n\bar{X} + m\bar{Y}}{n+m}$ is a better estimate since variance is smaller.

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 6**

**Question:**

A machine is set to fill bags with flour such that the mean weight is 1010 grams. To check that the machine is working properly, a random sample of 8 bags is selected. The weight of flour, in grams, in each bag is as follows.

1010 1015 1005 1000 998 1008 1012 1007

Carry out a suitable test, at the 5% significance level, to test whether or not the mean weight of flour in the bags is less than 1010 grams. (You may assume that the weight of flour delivered by the machine is normally distributed.)

**Solution:**

Let $x$ represent weight of flour

$\sum x = 8055 \quad \therefore \bar{x} = 1006.875$

$\sum x^2 = 8\ 110\ 611 \therefore s^2 = \frac{1}{7}\left\{8\ 110\ 611 - \frac{8055^2}{8}\right\} = 33.26785\ldots$

$\therefore s = 5.767825$

$H_0 : \mu = 1010; H_1 : \mu < 1010$

critical value: $t = -1.895$ so critical region $t < -1.895$

$$t = \frac{(1006.875 - 1010)}{\left(\frac{5.7678}{\sqrt{8}}\right)} = -1.5324$$

Since $-1.53$ is not in the critical region there is insufficient evidence to reject $H_0$.

The mean weight of flour delivered by the machine is 1010g.

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 7**

**Question:**

A train company claims that the probability $p$ of one of its trains arriving late is 10%. A regular traveler on the company's trains believes that the probability is greater than 10% and decides to test this by randomly selecting 12 trains and recording the number, $X$, of trains that were late. The traveller sets up the hypotheses $H_0: p = 0.1$ and $H_1: p > 0.1$ and accepts the null hypothesis if $x \leq 2$.

**a** Find the size of the test.

**b** Show that the power function of the test is $1-(1-p)^{10}(1+10p+55p^2)$.

**c** Calculate the power of the test when

**i** $p = 0.2$

**ii** $p = 0.6$

**d** Comment on your results from part **c**. **[E]**

**Solution:**

**a** $1-0.8891 = 0.1109$

**b** $1-(P(0)+P(1)+P(2))$

$= 1-((1-p)^{12}+12p(1-p)^{11}+66p^2(1-p)^{10})$

$= 1-(1-p)^{10}((1-p)^2+12p(1-p)+66p^2)$

$= 1-(1-p)^{10}\left[1-2p+p^2+12p-12p^2+66p^2\right]$

$= 1-(1-p)^{10}(1+10p+55p^2)$

**c** $1-0.5583 = 0.442$

$1-0.00281 = 0.997$

**d** The test is more discriminating (powerful) for the larger value of $p$.

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 8**

**Question:**

It is suggested that a Poisson distribution with parameter $\lambda$ can model the number of currants in a currant bun. A random bun is selected in order to test the hypotheses $H_0: \lambda = 8$ against $H_1: \lambda \neq 8$, using a 10% level of significance.

**a** Find the critical region for this test, such that the probability in each tail is as close as possible to 5%.

**b** Given that $\lambda = 10$, find

**i** the probability of a type II error,

**ii** the power of the test. **[E]**

**Solution:**

**a** $P(X \le c_1) \le 0.05; P(X \le 3 \mid \lambda = 8) = 0.0424 \Rightarrow X \le 3$

$P(X \ge c_2) \le 0.05; P(X \ge 14 \mid \lambda = 8) = 0.0342$

$P(X \ge 13 \mid \lambda = 8) = 0.0638 \Rightarrow X \ge 13$

$\therefore$ critical region is $\{X \le 3\} \cup \{X \ge 13\}$

**b** **i** $P(4 \le X \le 12 \mid \lambda = 10) = P(X \le 12) - P(X \le 3)$

$= 0.7916 - 0.0103$

$= 0.7813$

**ii** $\text{Power} = 1 - 0.7813 = 0.2187$

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 9**

**Question:**

The length $X$ mm of a spring made by a machine is normally distributed $N(\mu, \sigma^2)$.
A random sample of 20 springs is selected and their lengths measured in millimetres.
Using this sample, the unbiased estimates of $\mu$ and $\sigma^2$ are $\bar{x} = 100.6$ $s^2 = 1.5$
Stating your hypotheses clearly test, at the 10% level of significance,

**a** whether or not the variance of the lengths of springs is different from 0.9,

**b** whether or not the mean length of the springs is greater than 100 mm. **[E]**

**Solution:**

**a** $H_0 : \sigma^2 = 0.9$ $H_1 : \sigma^2 \neq 0.9$

$\nu = 19$

CR (Lower tail 10.117)

Upper tail 30.144

Test statistic $= \dfrac{19 \times 1.5}{0.9} = 31.6666$, significant

There is sufficient evidence that the *variance* of the length of spring is *different from 0.9*

**b** $H_0 : \mu = 100$ $H_1 : \mu > 100$

$t_{19} = 1.328$ is the critical value

$$t = \frac{100.6 - 100}{\sqrt{\frac{1.5}{20}}} = 2.19$$

Significant. The mean *length of spring* is *greater than 100*

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 10**

**Question:**

A town council is concerned that the mean price of renting two bedroom flats in the town has exceeded £650 per month. A random sample of eight two bedroom flats gave the following results, £$x$, per month.

705, 640, 560, 680, 800, 620, 580, 760

[You may assume $\sum x = 5345, \sum x^2 = 3\,621\,025$.]

**a** Find a 90% confidence interval for the mean price of renting a two bedroom flat.

**b** State an assumption that is required for the validity of your interval in part **a**.

**c** Comment on whether or not the town council is justified in being concerned.

Give a reason for your answer. **[E]**

**Solution:**

**a** $\bar{x} = 668.125 \quad s = 84.425$

$t_7(5\%) = 1.895$

$$\text{Confidence limits} = 668.125 \pm \frac{1.895 \times 84.425}{\sqrt{8}} = 611.6 \text{ and } 724.7$$

Confidence interval = (612, 725)

**b** Normal distribution

**c** £650 is within the confidence interval. No need to worry.

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 11**

**Question:**

A technician is trying to estimate the area $\mu^2$ of a metal square. The independent random variables $X_1$ and $X_2$ are each distributed $\mathrm{N}(\mu, \sigma^2)$ and represent two measurements of the sides of the square. Two estimators of the area, $A_1$ and $A_2$, are proposed where $A_1 = X_1X_2$ and $A_2 = \left(\frac{X_1 + X_2}{2}\right)^2$.

[You may assume that if $X_1$ and $X_2$ are independent random variables then $\mathrm{E}(X_1X_2) = \mathrm{E}(X_1)\mathrm{E}(X_2)$]

**a** Find $\mathrm{E}(A_1)$ and show that $\mathrm{E}(A_2) = \mu^2 + \frac{\sigma^2}{2}$.

**b** Find the bias of each of these estimators.

The technician is told that $\mathrm{Var}(A_1) = \sigma^4 + 2\mu^2\sigma^2$ and $\mathrm{Var}(A_2) = \frac{1}{2}\sigma^4 + 2\mu^2\sigma^2$.

The technician decided to use $A_1$ as the estimator for $\mu^2$.

**c** Suggest a possible reason for this decision.

A statistician suggests taking a random sample of $n$ measurements of sides of the square and finding the mean $\bar{X}$. He knows that $\mathrm{E}(\bar{X}^2) = \mu^2 + \frac{\sigma^2}{n}$ and $\mathrm{Var}(\bar{X}^2) = \frac{2\sigma^4}{n^2} + \frac{4\sigma^2\mu^2}{n}$.

**d** Explain whether or not $\bar{X}^2$ is a consistent estimator of $\mu^2$. **[E]**

**Solution:**

**a** $E(A_1) = E(X_1)E(X_2) = \mu^2$

$A_2 = \bar{X}^2, \bar{X} \sim N\left(\mu, \frac{\sigma^2}{2}\right) \therefore E(\bar{X}^2) = E(A_2) = \mu^2 + \frac{\sigma^2}{2}$

alternative for $A_2$

$E(A_2) = E\left[\left(\frac{X_1 + X_2}{2}\right)^2\right]$

$= E\left[\frac{X_1^2 + 2X_1X_2 + X_2^2}{4}\right]$

$= E\left[\frac{X_1^2}{4}\right] + E\left[\frac{X_1X_2}{2}\right] + E\left[\frac{X_2^2}{4}\right]$

$= \frac{1}{4}E\left(X_1^2\right) + \frac{1}{2}E\left(X_1X_2\right) + \frac{1}{4}E\left(X_2^2\right)$

but $Var(X) = E(X^2) - \mu^2$

so $\mu^2 + Var X = E(X^2)$

$\therefore E\left(X_1\right)^2 = \mu^2 + \sigma^2$ and $E\left(X_2^2\right) = \mu^2 + \sigma^2$

$\therefore E\left(A_2\right) = \frac{1}{4}\left(\mu^2 + \sigma^2\right) + \frac{1}{2}\mu^2 + \frac{1}{4}\left(\mu^2 + \sigma^2\right)$

$= \mu^2 + \frac{\sigma^2}{2}$

**b** $A_1$ is unbiased, bias for $A_2$ is $\frac{\sigma^2}{2}$

**c** Used $A_1$ since it is unbiased

**d** $E(\bar{X}^2) = \mu^2 + \frac{\sigma^2}{n}$; as $n \to \infty, E(\bar{X}^2) \to \mu^2$

$Var(\bar{X}^2) = \frac{2\sigma^4}{n^2} + \frac{4\sigma^2\mu^2}{n}$; as $n \to \infty, Var(\bar{X}^2) \to 0$

$\bar{X}^2$ is a consistent estimator of $\mu^2$

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 12**

**Question:**

A random sample of 15 tomatoes is taken and the weight $x$ grams of each tomato is found. The results are summarised by $\sum x = 208$ and $\sum x^2 = 2962$.

**a** Assuming that the weights of the tomatoes are normally distributed, calculate the 90% confidence interval for the variance $\sigma^2$ of the weights of the tomatoes.

**b** State, with a reason, whether or not the confidence interval supports the assertion $\sigma^2 = 3$. **[E]**

**Solution:**

**a** $s^2 = \dfrac{2962 - 15 \times \left(\dfrac{208}{15}\right)^2}{14} = 5.55$

$\dfrac{14 \times 5.55}{23.685} < \sigma^2 < \dfrac{14 \times 5.55}{6.571}$

$3.28 < \sigma^2 < 11.83$

**b** Since 9 lies in the interval, yes, it supports the assertion.

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 13**

**Question:**

**a** Define
**i** a type I error,
**ii** a type II error.

A small aviary, that leaves the eggs with the parent birds, rears chicks at an average rate of 5 per year. In order to increase the number of chicks reared per year it is decided to remove the eggs from the aviary as soon as they are laid and put them in an incubator. At the end of the first year of using an incubator 7 chicks had been successfully reared.

**b** Assuming that the number of chicks reared per year follows a Poisson distribution test, at the 5% significance level, whether or not there is evidence of an increase in the number of chicks reared per year. State your hypotheses clearly.

**c** Calculate the probability of the type I error for this test.

**d** Given that the true average number of chicks reared per year when the eggs are hatched in an incubator is 8, calculate the probability of a type II error. **[E]**

**Solution:**

**a** Type I – $H_0$ rejected when it is true

Type II – $H_0$ is accepted when it is false

**b** $H_0: \lambda = 5, H_1: \lambda > 5$

$P(X \geq 7 \mid \lambda = 5) = 1 - 0.7622 = 0.2378 > 0.05$

No evidence of an increase in the number of chicks reared per year.

**c** $P(X \geq c \mid \lambda = 5) < 0.05$

$P(X \geq 9) = 0.0681, P(X \geq 10) = 0.0318, c = 10$

$P(\text{Type I Error}) = 0.0318$

**d** $\lambda = 8$

$P(X \leq 9 \mid \lambda = 8) = 0.7166$

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 14**

**Question:**

**a** Explain briefly what you understand by
**i** an unbiased estimator,
**ii** a consistent estimator
of an unknown population parameter $\theta$.

From a binomial population, in which the proportion of successes is $p$, 3 samples of size $n$ are taken. The number of successes $X_1, X_2$, and $X_3$ are recorded and used to estimate $p$.

**b** Determine the bias, if any, of each of the following estimators of $p$.

$$\hat{p}_1 = \frac{X_1 + X_2 + X_3}{3n},$$
$$\hat{p}_2 = \frac{X_1 + 3X_2 + X_3}{6n},$$
$$\hat{p}_3 = \frac{2X_1 + 3X_2 + X_3}{6n}.$$

**c** Find the variance of each of these estimators.

**d** State, giving a reason, which of the three estimators for $p$ is
**i** the best estimator,
**ii** the worst estimator. **[E]**

**Solution:**

**a i** $\mathrm{E}(\hat{\theta}) = \theta$

**ii** $\mathrm{E}(\hat{\theta}) = \theta$ or $\mathrm{E}(\hat{\theta}) \to \theta$
and $\mathrm{Var}(\hat{\theta}) \to 0$ as $n \to \infty$ where $n$ is the sample size

**b** $\mathrm{E}(\hat{p}_1) = p$, $\therefore$ Bias $= 0$

$\mathrm{E}(\hat{p}_2) = \frac{5p}{6}$, $\therefore$ Bias $= -\frac{1}{6}p$

$\mathrm{E}(\hat{p}_3) = p$, $\therefore$ Bias $= 0$

**c** $\mathrm{Var}(\hat{p}_1) = \frac{1}{9n^2}\{npq + npq + npq\}$

$= \frac{pq}{3n}$ or $\frac{12pq}{36n}$

$\mathrm{Var}(\hat{p}_2) = \frac{1}{36n^2}\{npq + 9npq + npq\} = \frac{11pq}{36n}$

$\mathrm{Var}(\hat{p}_3) = \frac{1}{36n^2}\{4npq + 9npq + npq\} = \frac{7pq}{18n}$ or $\frac{14pq}{36n}$

**d i** $\hat{p}_1$; unbiased and smallest variance

**ii** $\hat{p}_2$; biased

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 15**

**Question:**

Define

**a** a type I error,

**b** the size of a test.

Jane claims that she can read Alan's mind. To test this claim Alan randomly chooses a card with one of 4 symbols on it. He then concentrates on the symbol. Jane then attempts to read Alan's mind by stating what symbol she thinks is on the card. The experiment is carried out 8 times and the number of times, $X$, that Jane is correct is recorded.

The probability of Jane stating the correct symbol is denoted by $p$.

To test the hypothesis $H_0 : p = 0.25$ against $H_1 : p > 0.25$, a critical region of $X > 6$ is used.

**c** Find the size of this test.

**d** Show that the power function of this test is $8p^7 - 7p^8$.

Given that $p = 0.3$, calculate

**e** the power of this test,

**f** the probability of a type II error.

**g** Suggest two ways in which you might reduce the probability of a type II error. **[E]**

**Solution:**

**a** A Type I error occurs when $H_0$ is rejected when it is in fact true.

**b** The size of a test is the probability of a type I error.

**c** $X \sim B(8, 0.25)$

$$\text{Size} = P(X > 6) = 1 - P(X \le 6 \mid n = 8, p = 0.25)$$
$$= 1 - 0.9996 = 0.0004$$

**d** $\text{Power} = P(X > 6 \mid p = p, n = 8)$

$$= P(X = 7) + P(X = 8) = \frac{8}{7!1!} p^7 (1 - p) + p^8$$
$$= 8p^7 - 8p^8 + p^8 = 8p^7 - 7p^8$$

**e** $\text{Power} = 8 \times 0.3^7 - 7 \times 0.3^8$

$$= 0.00129$$

**f** $\text{Power} = 1 - P(\text{Type II})$

$$\therefore P(\text{Type II}) = 1 - \text{power}$$
$$= 1 - 0.00129$$
$$= 0.9987$$

**g** Increase the probability of a Type I error, e.g. increase the significance level of the test. Increase the value of Po

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 16**

**Question:**

Rolls of cloth delivered to a factory contain defects at an average rate of $\lambda$ per metre. A quality assurance manager selects a random sample of 15 metres of cloth from each delivery to test whether or not there is evidence that $\lambda > 0.3$. The criterion that the manager uses for rejecting the hypothesis that $\lambda = 0.3$ is that there are 9 or more defects in the sample.

**a** Find the size of the test.

Table 1 gives some values, to 2 decimal places, of the power function of this test.

| $\lambda$ | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1.0 |
|---|---|---|---|---|---|---|---|
| **Power** | 0.15 | 0.34 | $r$ | 0.72 | 0.85 | 0.92 | 0.96 |

Table 1

**b** Find the value of $r$.

The manager would like to design a test, of whether or not $\lambda > 0.3$, that uses a smaller length of cloth. He chooses a length of 10 m and requires the probability of a type I error to be less than 10%.

**c** Find the criterion to reject the hypothesis that $\lambda = 0.3$ which makes the test as powerful as possible.

**d** Hence state the size of this second test.

Table 2 gives some values, to 2 decimal places, of the power function for the test in part **c**.

| $\lambda$ | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1.0 |
|---|---|---|---|---|---|---|---|
| **Power** | 0.21 | 0.38 | 0.55 | 0.70 | $s$ | 0.88 | 0.93 |

Table 2

**e** Find the value of $s$.

**f** Using the same axes, on graph paper draw the graphs of the power functions of these two tests.

**g** **i** State the value of $\lambda$ where the graphs cross.

**ii** Explain the significance of $\lambda$ being greater than this value.

The cost of wrongly rejecting a delivery of cloth with $\lambda = 0.3$ is low. Deliveries of cloth with $\lambda > 0.7$ are unusual.

**h** Suggest, giving your reasons, which the test manager should adopt. **[E]**

**Solution:**

**a** $X_1 =$ no. of defects in 15 m $X_1 \sim \text{Po}(4.5)$
Size $= \text{P}(X_2 \geq 9) = 1 - 0.9597 = 0.0403$

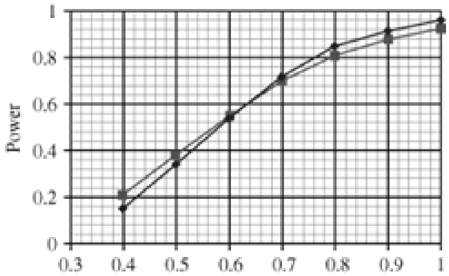
**b** $r = \text{P}(X_2 \geq 9 \mid X_2 \sim \text{Po}(9)) = 1 - 0.4557 = 0.54(43)$

**c** $Y_1 =$ No of defects in 10 m $Y_1 \sim \text{Po}(3)$
$\text{P}(Y_1 \geq c) < 0.10$ $Y_1 \geq 6$

**d** Size $= \text{P}(Y_1 \geq 6) = 1 - \text{P}(Y_1 \leq 5) = 1 - 0.9161 = 0.0839$

**e** $s = 1 - \text{P}(Y_2 \leq 5) = 1 - 0.1912 = 0.8088$

**f**




**g** **i** 0.62 to 0.67
**ii** Test I more powerful

**h** Test 2 Has a higher P(Type I error) but cost of this is low.
Test 2 is more powerful for $\lambda < 0.7$ and $\lambda > 0.7$ is rare.
Adopt test 2

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 17**

**Question:**

The number of tornadoes per year to hit a particular town follows a Poisson distribution with mean $\lambda$. A weatherman claims that due to climate changes the mean number of tornadoes per year has decreased. He records the number of tornadoes $x$ to hit the town last year.

To test the hypotheses $H_0 : \lambda = 7$ and $H_1 : \lambda < 7$, a critical region of $x \leq 3$ is used.

**a** Find, in terms $\lambda$ the power function of this test.

**b** Find the size of this test.

**c** Find the probability of a type II error when $\lambda = 4$. **[E]**

**Solution:**

**a** $\text{Power} = P(X \leq 3 \quad \lambda = 3)$

$$= e^{-\lambda} + e^{-\lambda}\lambda + \frac{e^{-\lambda}\lambda^2}{2} + \frac{e^{-\lambda}\lambda^3}{6}$$

$$= \frac{e^{-\lambda}}{6}(6 + 6\lambda + 3\lambda^2 + \lambda^3)$$

**b** CR is $X \leq 3$

$\text{Size} = P(X \leq 3 \quad \lambda = 7)$

$= 0.0818$

**c** $P(\text{Type II error}) = 1 - \text{power}$

$$= 1 - \frac{e^{-4}}{6}(6 + 6 \times 4 + 3 \times 4^2 + 4^3)$$

$$= 0.5665..$$

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 18**

**Question:**

A nutritionist studied the levels of cholesterol, $X$ mg/cm$^3$, of male students at a large college. She assumed that $X$ was distributed $N(\mu, \sigma^2)$ and examined a random sample of 25 male students. Using this sample she obtained unbiased estimates of $\mu$ and $\sigma^2$ as

$\hat{\mu} = 1.68 \quad \hat{\sigma}^2 = 1.79$

**a** Find a 95% confidence interval for $\mu$.

**b** Obtain a 95% confidence interval for $\sigma^2$.

A cholesterol reading of more than 2.5 mg/cm$^3$ is regarded as high.

**c** Use appropriate confidence limits from parts **a** and **b** to find the lowest estimate of the proportion of male students in the college with high cholesterol. **[E]**

**Solution:**

**a** 95% confidence interval for $\mu$ is

$$1.68 \pm t_{24}(2.5\%)\sqrt{\frac{1.79}{25}} = 1.68 \pm 2.064\sqrt{\frac{1.79}{25}} = (1.13, 2.23)$$

**b** 95% confidence interval for $\sigma^2$ is

$$12.401 < \frac{24 \times 1.79}{\sigma^2} < 39.364$$

$\sigma^2 > 1.09, \sigma^2 < 3.46$

$\therefore$ confidence interval on $\sigma^2$ is (1.09, 3.46)

**c** Require $P(X > 2.5) = P\left(Z > \frac{2.5 - \mu}{\sigma}\right)$ to be as small as possible OR

$\frac{2.5 - \mu}{\sigma}$ to be as large as possible; both imply lowest $\sigma$ and $\mu$.

$$\frac{2.5 - 1.13}{\sqrt{1.09}} = 1.31$$

$P(Z > 1.31) = 1 - 0.9049 = 0.0951$

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 19**

**Question:**

A random sample of three independent variables $X_1, X_2$ and $X_3$ is taken from a distribution with mean $\mu$ and variance $\sigma^2$.

**a** Show that $\frac{2}{3}X_1 - \frac{1}{2}X_2 + \frac{5}{6}X_3$ is an unbiased estimator for $\mu$.

An unbiased estimator for $\mu$ is given by $\hat{\mu} = aX_1 + bX_2$ where $a$ and $b$ are constants.

**b** Show that $\text{Var}(\hat{\mu}) = (2a^2 - 2a + 1)\sigma^2$.

**c** Hence determine the value of $a$ and the value of $b$ for which $\hat{\mu}$ has minimum variance. **[E]**

**Solution:**

**a** $\text{E}\left(\frac{2}{3}X_1 - \frac{1}{2}X_2 + \frac{5}{6}X_3\right) = \frac{2}{3}\mu - \frac{1}{2}\mu + \frac{5}{6}\mu = \mu$

$\text{E}(Y) = \mu \Rightarrow \text{unbiased}$

**b**
$$\begin{aligned} \text{E}(aX_1 + bX_2) &= a\mu + b\mu = \mu \\ a + b &= 1 \\ \text{Var}(aX_1 + bX_2) &= a^2\sigma^2 + b^2\sigma^2 \\ &= a^2\sigma^2 + (1-a)^2\sigma^2 \\ &= (2a^2 - 2a + 1)\sigma^2 \end{aligned}$$

**c** Minimum value when $(4a - 2)\sigma^2 = 0$ (from differentiation)

$\Rightarrow 4a - 2 = 0$

$a = \frac{1}{2}, b = \frac{1}{2}$

# Solutionbank S4

## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 20**

**Question:**

A supervisor wishes to check the typing speed of a new typist. On 10 randomly selected occasions, the supervisor records the time taken for the new typist to type 100 words. The results, in seconds, are given below.

| | | | | | | |
|---|---|---|---|---|---|---|
| 110 | 125 | 130 | 126 | 128 | 127 | 118 |
| 120 | 122 | 125 | | | | |

The supervisor assumes that the time taken to type 100 words is normally distributed.

**a** Calculate a 95% confidence interval for

**i** the mean,

**ii** the variance

of the population of times taken by this typist to type 100 words.

The supervisor requires the average time needed to type 100 words to be no more than 130 seconds and the standard deviation to be no more than 4 seconds.

**b** Comment on whether or not the supervisor should be concerned about the speed of the new typist. **[E]**

**Solution:**

**a** $\bar{x} = 123.1$

$s = 5.87745\ldots$

(NB: $\Sigma x = 1231$; $\Sigma x^2 = 151\ 847$)

**i** 95% confidence interval is given by

$$123.1 \pm 2.262 \times \frac{5.87745\ldots}{\sqrt{10}}$$

i.e. (118.8958 …, 127.30418 …)

**ii** 95% confidence interval is given by

$$\frac{9 \times 5.87745\ldots^2}{19.023} < \sigma^2 < \frac{9 \times 5.87745\ldots^2}{2.700}$$

i.e. (16.34336 …, 115.14806 ….)

**b** 130 is just above confidence interval

16 is just below confidence interval

Thus supervisor should be concerned about the speed of the new typist since both their average speed is two slow and the variability of the time is too large

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 21**

**Question:**

A machine is filling bottles of milk. A random sample of 16 bottles was taken and the volume of milk in each bottle was measured and recorded. The volume of milk in a bottle is normally distributed and the unbiased estimate of the variance, $s^2$, of the volume of milk in a bottle is 0.003.

**a** Find a 95% confidence interval for the variance of the population of volumes of milk from which the sample was taken.

The machine should fill bottles so that the standard deviation of the volumes is equal to 0.07.

**b** Comment on this with reference to your 95% confidence interval. **[E]**

**Solution:**

**a** $\text{Confidence interval} = \left(\frac{15\times 0.003}{27.488}, \frac{15\times 0.003}{6.262}\right)$

$= (0.00164, 0.00719)$

**b** $0.07^2 = 0.0049$

0.0049 is within the 95% confidence interval.

There is no evidence to reject the idea that the standard deviation of the volumes is not 0.07 or the machine is working well.

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 22**

**Question:**

A butter packing machine cuts butter into blocks. The weight of a block of butter is normally distributed with a mean weight of 250 g and a standard deviation of 4 g.
A random sample of 15 blocks is taken to monitor any change in the mean weight of the blocks of butter.

**a** Find the critical region of a suitable test using a 2% level of significance.

**b** Assuming the mean weight of a block of butter has increased to 254 g, find the probability of a type II error. **[E]**

**Solution:**

**a** $\frac{\bar{X}-250}{\frac{4}{\sqrt{15}}} > 2.3263$ or $\frac{\bar{X}-250}{\frac{4}{\sqrt{15}}} < -2.3263$

$\bar{X} > 252.40\ldots$ or $\bar{X} < 247.6\ldots$

**b** $\mathrm{P}(\bar{X} > 252.4 / \mu = 254) - \mathrm{P}(\bar{X} < 247.6 / \mu = 254)$

$= \mathrm{P}\left(Z > \frac{252.4-254}{\frac{4}{\sqrt{15}}}\right) - \mathrm{P}\left(Z < \frac{247.6-254}{\frac{4}{\sqrt{15}}}\right)$

$= \mathrm{P}(Z > -1.5492) - \mathrm{P}(Z < -6.20)$

$= (1-0.9394)-(0)$

$= 0.0606$

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 23**

**Question:**

A drug is claimed to produce a cure to a certain disease in 35% of people who have the disease. To test this claim a sample of 20 people having this disease is chosen at random and given the drug. If the number of people cured is between 4 and 10 inclusive the claim will be accepted.

**a** Write down suitable hypotheses to carry out this test.

**b** Find the probability of making a type I error.

The table below gives the value of the probability of the type II error, to 4 decimal places, for different values of $p$ where $p$ is the probability of the drug curing a person with the disease.

| **P(cure)** | 0.2 | 0.3 | 0.4 | 0.5 |
|---|---|---|---|---|
| **P(Type II error)** | 0.5880 | $r$ | 0.8565 | $s$ |

**c** Calculate the value of $r$ and the value of $s$.

**d** Calculate the power of the test for $p = 0.2$ and $p = 0.4$

**e** Comment, giving your reasons, on the suitability of this test procedure. **[E]**

**Solution:**

**a** $H_0 : p = 0.35 \quad H_1 : p \neq 0.35$

**b** Let $X =$ Number cured then $X \sim B(20, 0.35)$

$\alpha = P(\text{Type I error}) = P(x \leq 3) + P(x \geq 11)$ given $p = 0.35$

$= 0.0444 + 0.0532$

$= 0.0976$

**c** $\beta = P(\text{Type II error}) = P(4 \leq \times \leq 10)$

| $p$ | 0.2 | 0.3 | 0.4 | 0.5 |
|---|---|---|---|---|
| $\beta$ | 0.5880 | 0.8758 | 0.8565 | 0.5868 |

**d** Power $= 1 - \beta$

0.4120 0.1435

**e** Not a good procedure.
Better further away from 0.35 or
This is not a very powerful test (power $= 1 - \beta$)

# Solutionbank S4
## Edexcel AS and A Level Modular Mathematics

**Review Exercise 1**
**Exercise A, Question 24**

**Question:**

A doctor wishes to study the level of blood glucose in males. The level of blood glucose is normally distributed. The doctor measured the blood glucose of 10 randomly selected male students from a school. The results, in mmol/litre, are given below.

| | | | | |
|---|---|---|---|---|
| 4.7 | 3.6 | 3.8 | 4.7 | 4.1 |
| 2.2 | 3.6 | 4.0 | 4.4 | 5.0 |

**a** Calculate a 95% confidence interval for the mean.

**b** Calculate a 95% confidence interval for the variance.

A blood glucose reading of more than 7 mmol/litre is counted as high.

**c** Use appropriate confidence limits from parts **a** and **b** to find the highest estimate of the proportion of male students in the school with a high blood glucose level. **[E]**

**Solution:**

$\bar{x} = 4.01$

$s = 0.7992\ldots$

**a** $4.01 \pm t_9(2.5\%)\dfrac{0.7992..}{\sqrt{10}} = 4.01 \pm 2.262\dfrac{0.7992..}{\sqrt{10}}$

$= 4.5816\ldots$ and $3.4383\ldots$

i.e. (3.4383, 4.5816.)

**b** $2.700 < \dfrac{9 \times 0.7992\ldots^2}{s^2} < 19.023$

$\sigma^2 < 2.13, \sigma^2 > 0.302$

i.e. (0.302, 2.13)

**c** $\mathrm{P}(X > 7) = \mathrm{P}\left(Z > \dfrac{7-\mu}{\sigma}\right)$ needs to be as high as possible

Therefore $\mu$ and $\sigma$ must be as big as possible

propertion with high blood glucose level $= \mathrm{P}\left(Z > \dfrac{7-4.581}{\sqrt{2.13}}\right)$

$= 1 - 0.9515$

$= 0.0485$

$= 4.85\%$