# IB Math AI HL — Topic 4

## Normal Distributions, Sample Means & Confidence Intervals

Combinations of Normals · Sample Mean Distribution · Central Limit Theorem · Confidence Interval for the Mean

| Difficulty | EASY | MEDIUM | HARD |
|---|---|---|---|
| Questions | 10 | 10 | 10 |
| Total Marks | 30 | 40 | 50 |

Scan me for help

Name:____________________ Date:__________

Score:________/120

**Instructions:** Answer all questions. Show all working. Give answers correct to 3 significant figures unless stated. GDC permitted throughout.

### Key Formulae

**Normal Distribution:** $X \sim \mathrm{N}(\mu, \sigma^2)$

**Standardising:** $Z = \dfrac{X-\mu}{\sigma}, \quad Z \sim \mathrm{N}(0,1)$

**Linear Combination:**
If $X \sim \mathrm{N}(\mu_X, \sigma_X^2)$ and $Y \sim \mathrm{N}(\mu_Y, \sigma_Y^2)$, independent:
$aX + bY \sim \mathrm{N}(a\mu_X + b\mu_Y,\ a^2\sigma_X^2 + b^2\sigma_Y^2)$

**Sum of $n$ identical copies:**
$X_1 + X_2 + \cdots + X_n \sim \mathrm{N}(n\mu,\ n\sigma^2)$

**Sample Mean Distribution:**
$\bar{X} \sim \mathrm{N}\left(\mu, \dfrac{\sigma^2}{n}\right), \quad \mathrm{SE} = \dfrac{\sigma}{\sqrt{n}}$

**Central Limit Theorem (CLT):**
For large $n$ (typically $n \geq 30$), $\bar{X}$ is approximately normal regardless of the population distribution.

**95% Confidence Interval for $\mu$:**
$\bar{x} \pm z^* \cdot \dfrac{s_{n-1}}{\sqrt{n}}, \quad z^* = 1.96$ (95%)

**Unbiased estimate of $\sigma^2$:**
$s_{n-1}^2 = \dfrac{n}{n-1}\, s_n^2$

### GDC Tips

**Normal probabilities:** Use `normalcdf(lower, upper, μ, σ)` for $P(a < X < b)$; use $-1 \times 10^{99}$ for $-\infty$ and $1 \times 10^{99}$ for $+\infty$.

**Inverse normal:** Use `invNorm(area, μ, σ)` for the LEFT-tail area.

**Confidence intervals:** `STAT` → `TESTS` → `TInterval` (use $s_{n-1}$). Enter $\bar{x}$, $s_{n-1}$, $n$, confidence level.

**Sample mean distribution:** Replace $\sigma$ with $\sigma/\sqrt{n}$ in all normal GDC commands.

**Common Mistakes to Avoid**

1. **SD vs Variance in combinations.** Variances add, NOT standard deviations: Var$(X+Y) = \sigma_X^2 + \sigma_Y^2$. Never write $\sigma_{X+Y} = \sigma_X + \sigma_Y$.
2. **Sample mean SE.** The standard error of $\bar{X}$ is $\sigma/\sqrt{n}$, not $\sigma/n$ or $\sigma \cdot \sqrt{n}$.
3. $s_n$ **vs** $s_{n-1}$**.** Use $s_{n-1}$ (unbiased) for confidence intervals and as an estimator of $\sigma$. Your GDC gives both — use $S_x$ (not $\sigma_x$).
4. **Scaling a normal.** $2X \neq X + X'$ (where $X'$ is independent). Var$(2X) = 4\sigma^2$, while Var$(X + X') = 2\sigma^2$.
5. **CI interpretation.** A 95% CI does NOT mean "95% probability the true mean is in this interval." It means that 95% of such intervals constructed from repeated samples would contain the true mean.
6. **CLT conditions.** State that $n$ is large (typically $\geq 30$) when invoking CLT for non-normal populations.

## Section A — Easy EASY (10 questions × 3 marks = 30 marks)

**Scan me for help**

**1.** EASY GDC **[3 marks]**

The daily output of a bottling machine can be modelled by a normal distribution with mean $\mu = 4800$ bottles and standard deviation $\sigma = 120$ bottles.

(a) Find the probability that on a randomly chosen day the machine produces more than $4950$ bottles. [2]

(b) Find the number of bottles $k$ such that $P(\text{output} < k) = 0.10$. [1]

**2.** EASY GDC **[3 marks]**

The resting heart rate of adult volunteers in a health study is modelled by $H \sim \text{N}(72, 8^2)$, where $H$ is measured in beats per minute.

(a) Write down the mean and standard deviation of $H$. [1]

(b) Find $P(65 < H < 85)$. [2]

**3.** EASY GDC **[3 marks]**

The mass of bags of dried pasta produced by a factory follows $M \sim \text{N}(500, 6^2)$, where $M$ is in grams. A quality inspector selects a random sample of $n = 36$ bags.

(a) Write down the distribution of the sample mean $\bar{M}$, stating its mean and variance. [2]

(b) Find the standard error of $\bar{M}$. [1]

**4.** EASY GDC **[3 marks]**

Two independent random variables are defined as $X \sim \text{N}(30, 4^2)$ and $Y \sim \text{N}(20, 3^2)$.

(a) Find the mean of $W = X + Y$. [1]

(b) Find the variance of $W = X + Y$. [1]

(c) Write down the distribution of $W$. [1]

**5.** EASY GDC **[3 marks]**

A random sample of $n = 25$ jars of honey is selected from a production line. The sample mean net weight is $\bar{x} = 453$ g and the sample standard deviation is $s_{n-1} = 10$ g. Assume weights are normally distributed.

(a) Find a 95% confidence interval for the population mean net weight. [2]

(b) State one assumption needed for the interval to be valid. [1]

**6.** EASY GDC **[3 marks]**

The time taken by students to complete an online quiz follows a distribution with mean $\mu = 18$ minutes and standard deviation $\sigma = 4$ minutes. A random sample of $n = 64$ students is selected.

The table below summarises the distribution of the sample mean $\bar{T}$.

| **Parameter** | **Value** |
|---|---|
| Mean of $\bar{T}$ | $18$ min |
| Standard deviation of $\bar{T}$ | $0.5$ min |
| Approximate distribution | Normal (by CLT) |

(a) State the condition required to apply the Central Limit Theorem here. [1]

(b) Find $P(\bar{T} > 19)$. [2]

**7.** EASY GDC **[3 marks]**

The lengths of adult trout in a river follow $L \sim \mathrm{N}(38, 5^2)$, where $L$ is measured in centimetres.

(a) Find the value of $c$ such that $P(L > c) = 0.25$. [2]

(b) Find the probability that a randomly selected trout has a length between $30$ cm and $45$ cm. [1]

**8.** EASY GDC **[3 marks]**

A researcher calculates a 95% confidence interval for the mean daily step count of office workers. The interval is $(7\,240, 8\,360)$ steps.

(a) Write down the sample mean $\bar{x}$. [1]

(b) Write down the margin of error for this confidence interval. [1]

(c) Determine whether it is plausible that the true mean daily step count is $7\,000$ steps. Give a reason. [1]

**9.** EASY GDC **[3 marks]**

The fuel consumption of a certain model of car follows $C \sim \text{N}(7.2,\ 0.4^2)$ litres per $100$ km. A fleet manager selects a random sample of $16$ cars.

(a) State the distribution of the sample mean consumption $\bar{C}$. [1]

(b) Find $P(\bar{C} < 7.1)$. [2]

**10.** EASY GDC **[3 marks]**

The morning delivery time for a courier service follows $A \sim \text{N}(25, 4^2)$ minutes, and the afternoon delivery time follows $B \sim \text{N}(35, 5^2)$ minutes, independently.

(a) Find the mean and standard deviation of the total daily delivery time $T = A + B$. [2]

(b) Find $P(T > 70)$. [1]

## Section B — Medium **MEDIUM** (10 questions × 4 marks = 40 marks)

**Scan me for help**

**11.** MEDIUM GDC **[4 marks]**

A shop sells two types of flour. Type A has mass $A \sim \text{N}(1000, 15^2)$ grams and Type B has mass $B \sim \text{N}(500, 10^2)$ grams per bag. A customer buys two bags of Type A and three bags of Type B, independently.

Let $W = 2A + 3B$ be the total mass purchased.

(a) Find the mean of $W$. [1]

(b) Find the standard deviation of $W$. [2]

(c) Find the probability that the total mass exceeds $3\,550$ g. [1]

**12.** MEDIUM GDC **[4 marks]**

A marine biologist measures the shell diameter (in mm) of $n = 12$ randomly selected sea urchins. The data are shown below.

| **Shell diameter (mm)** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 43.2 | 47.8 | 45.1 | 44.6 | 46.3 | 48.9 | 42.7 | 45.5 | 47.2 | 44.0 | 46.8 | 43.9 |

(a) Calculate the unbiased estimates of the population mean and variance. [2]

(b) Hence find a 95% confidence interval for the true mean shell diameter. [2]

**13.** MEDIUM GDC **[4 marks]**

The mass of mangoes harvested from an orchard follows a normal distribution. It is known that $P(M < 180) = 0.20$ and $P(M > 260) = 0.15$, where $M$ is the mass in grams.

(a) Write down two equations in $\mu$ and $\sigma$. [2]

(b) Hence find $\mu$ and $\sigma$. [2]

**14.** MEDIUM GDC **[4 marks]**

A factory produces steel rods whose lengths are not normally distributed, but have population mean $\mu = 85$ cm and population standard deviation $\sigma = 3$ cm. A random sample of $n = 50$ rods is selected.

(a) State the approximate distribution of the sample mean length $\bar{L}$, justifying your answer. [2]

(b) Find $P(84.2 < \bar{L} < 85.6)$. [2]

**15.** MEDIUM GDC **[4 marks]**

A psychologist wishes to estimate the mean reaction time $\mu$ (in milliseconds) of adults to a visual stimulus. From previous studies, the population standard deviation is known to be $\sigma = 35$ ms. The psychologist wants the 95% confidence interval to have a margin of error of at most $8$ ms.

(a) Write down the formula for the margin of error of a 95% confidence interval. [1]

(b) Find the minimum sample size $n$ required. [3]

**16.** MEDIUM GDC **[4 marks]**

The heights of male students at a university follow $M \sim \mathrm{N}(178, 7^2)$ cm, and the heights of female students follow $F \sim \mathrm{N}(165, 6^2)$ cm, independently. A male and a female student are selected at random. Let $D = M - F$.

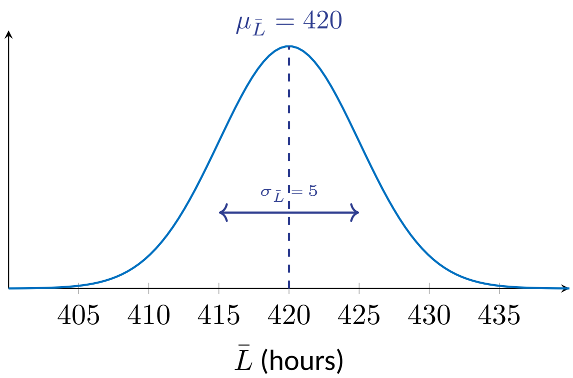
(a) Find the mean and standard deviation of $D$. [2]

(b) Find the probability that the male student is at least $20$ cm taller than the female student. [2]

**17.** MEDIUM GDC **[4 marks]**

The lifetime of a particular brand of rechargeable battery follows $L \sim \mathrm{N}(420, 25^2)$ hours. A random sample of $n = 25$ batteries is tested.

The diagram below shows the distribution of the sample mean lifetime $\bar{L}$.




(a) Write down the standard deviation of $\bar{L}$. [1]

(b) Find $P(\bar{L} > 425)$. [1]

(c) Find the value $k$ such that $P(\bar{L} > k) = 0.05$. [2]

**18.** MEDIUM GDC **[4 marks]**

A quality control manager selects $n = 20$ boxes of breakfast cereal. The net weights (in grams) give a sample mean of $\bar{x} = 497.4$ g and an unbiased sample standard deviation of $s_{n-1} = 8.2$ g. The population is assumed to be normally distributed.

(a) Find a 95% confidence interval for the true mean net weight. Give the bounds to 1 decimal place. [2]

(b) The box label states the net weight is $500$ g. Using your confidence interval, comment on whether this claim is plausible. [2]

**19.** MEDIUM GDC **[4 marks]**

Two independent machines produce metal washers. Machine X produces washers with diameter $X \sim \mathrm{N}(24.0, 0.3^2)$ mm, and Machine Y produces washers with diameter $Y \sim \mathrm{N}(24.0, 0.5^2)$ mm.

A combined washer assembly is formed by stacking one washer from Machine X and two washers from Machine Y. Let $S = X + 2Y$ represent the total stack thickness.

(a) Find the mean and standard deviation of $S$. [2]

(b) Find the probability that the total stack thickness is less than $71$ mm. [2]

**20.** MEDIUM GDC **[4 marks]**

Two brands of batteries are tested for their lifetime. The results are summarised in the table below.

| **Brand** | **Sample size $n$** | **Sample mean $\bar{x}$ (hours)** | **$s_{n-1}$ (hours)** |
|---|---|---|---|
| Volta | 30 | 18.4 | 2.1 |
| MaxPower | 30 | 17.9 | 1.8 |

(a) Find a 95% confidence interval for the mean lifetime of Volta batteries. [2]

(b) Find a 95% confidence interval for the mean lifetime of MaxPower batteries. [2]

## Section C — Hard **HARD** (10 questions × 5 marks = 50 marks)

**Scan me for help**

**21.** HARD GDC **[5 marks]**

A jeweller works with two types of gold chain. Chain P has mass $P \sim \mathrm{N}(12.0, 0.8^2)$ grams per link, and Chain Q has mass $Q \sim \mathrm{N}(8.0, 0.5^2)$ grams per link, independently. A necklace is assembled using 3 links of Chain P and 4 links of Chain Q.

Let $T$ be the total mass of the necklace.

(a) Show that $T \sim \mathrm{N}(68, 9.76)$. [3]

(b) Find the value of $m$ such that $P(T > m) = 0.10$. [2]

**22.** HARD GDC **[5 marks]**

The time taken (in minutes) for a technician to service a photocopier follows a normal distribution. It is known that $P(T < 45) = 0.30$ and $P(T > 75) = 0.10$.

(a) Find the mean $\mu$ and standard deviation $\sigma$ of the service time distribution. [3]

(b) A random sample of $n = 16$ service calls is observed. Find the probability that the sample mean service time exceeds $65$ minutes. [2]

**23.** HARD GDC **[5 marks]**

A dietitian is estimating the mean daily calcium intake (in mg) of adults aged 30–50. From pilot data, the population standard deviation is estimated as $\sigma = 180$ mg.

(a) Find the minimum sample size required so that a 99% confidence interval has a total width of at most $60$ mg. Use $z^* = 2.576$ for 99% confidence. [3]

(b) The dietitian collects data from $n = 250$ adults and obtains $\bar{x} = 1\,045$ mg. Find the 99% confidence interval for the mean daily calcium intake. [2]

**24.** HARD GDC **[5 marks]**

A food technologist blends two ingredients. Ingredient A has protein content $A \sim \mathrm{N}(22, 2^2)$ g per $100$ g, and Ingredient B has protein content $B \sim \mathrm{N}(15, 1.5^2)$ g per $100$ g. A blend consists of $60\%$ of Ingredient A and $40\%$ of Ingredient B (by mass).

The protein content of the blend per $100$ g is modelled by $W = 0.6A + 0.4B$.

(a) Find the mean and variance of $W$. [3]

(b) The blend meets a nutritional standard if $W > 19$ g per $100$ g. Find the probability that a randomly selected batch of blend meets this standard. [2]

**25.** HARD GDC **[5 marks]**

A botanist measures the stem length (in cm) of $n$ randomly selected seedlings of a new variety. The population is known to be normally distributed. The sample gives $\bar{x} = 14.3$ cm and $s_{n-1} = 2.4$ cm.

(a) Find the 95% confidence interval for the population mean stem length when $n = 16$. [2]

(b) Find the minimum value of $n$ such that the width of the 95% confidence interval is less than $1.5$ cm, assuming $s_{n-1}$ remains $2.4$ cm. [3]

**26.** HARD GDC **[5 marks]**

The mass of individual apples in a large orchard follows a distribution with mean $\mu = 185$ g and standard deviation $\sigma = 20$ g (not necessarily normal). Apples are packed into boxes of $n = 40$.

(a) State, with justification, the approximate distribution of the mean apple mass $\bar{M}$ per box. [2]

(b) A box is rejected if the mean apple mass per box is below $180$ g. Find the probability that a randomly selected box is rejected. [2]

(c) Find the value of $\mu_0$ such that $P(\bar{M} < 180) = 0.01$. [1]

**27.** HARD GDC **[5 marks]**

Two independent random variables are defined as $X \sim \mathrm{N}(\mu, 5^2)$ and $Y \sim \mathrm{N}(10, 3^2)$. Let $D = X - Y$.
It is known that $P(D > 8) = 0.20$.

(a) Find the mean of $D$ in terms of $\mu$. [1]

(b) Show that the standard deviation of $D$ is $\sqrt{34}$. [2]

(c) Hence find the value of $\mu$. [2]

**28.** HARD GDC **[5 marks]**

A logistics company tests two brands of van tyre. The table below summarises the tread-wear data (in thousands of km) from independent random samples, assuming normally distributed populations.

| **Brand** | $n$ | $\bar{x}$ **(000 km)** | $s_{n-1}$ **(000 km)** |
|---|---|---|---|
| ArcoTyre | 18 | 62.4 | 4.8 |
| RoadKing | 22 | 58.9 | 5.6 |

(a) Find a 95% confidence interval for the mean tread wear of ArcoTyre. [2]

(b) Find a 95% confidence interval for the mean tread wear of RoadKing. [2]

(c) Comment on whether the data provide evidence that the two brands have different mean tread wear. [1]

**29.** HARD GDC **[5 marks]**

The volume of liquid dispensed by a vending machine per cup follows $V \sim \mathrm{N}(200, \sigma^2)$ ml. Quality control requires that when $n = 25$ cups are sampled, the probability that the sample mean exceeds $202$ ml is at most $0.05$.

(a) Write down the distribution of $\bar{V}$ in terms of $\sigma$. [1]

(b) Show that the condition $P(\bar{V} > 202) \leq 0.05$ requires $\sigma \leq \dfrac{10}{1.645}$. [2]

(c) Hence find the maximum allowable value of $\sigma$, correct to 3 significant figures. [1]

(d) Given $\sigma = 6.0$ ml, find $P(198 < \bar{V} < 203)$. [1]

**30.** HARD GDC **[5 marks]**

A study examines the total daily calorie intake of adults. Breakfast calories follow $B \sim \text{N}(480, 60^2)$ kcal and dinner calories follow $D \sim \text{N}(720, 90^2)$ kcal, independently.

Let $C = B + D$ represent the combined breakfast and dinner calorie intake.

(a) Find the mean and standard deviation of $C$. [2]

(b) A nutritionist samples $n = 36$ adults and records each person's value of $C$. Find the probability that the sample mean $\bar{C}$ lies between $1\,170$ kcal and $1\,230$ kcal. [2]

(c) State the result that allows you to treat $\bar{C}$ as approximately normal in part (b), and identify the condition that must hold. [1]

## Worked Solutions — Official Mark Scheme

M = Method mark   A = Accuracy/Answer mark   R = Reasoning mark   AG = Answer Given (no mark)   (M1)/(A1) = implied mark

### Section A — Easy

#### Q1 — Normal Distribution: Basic Probability EASY

**(a)** $X \sim \text{N}(4800, 120^2)$

$P(X > 4950)$ **(M1)**

$= 1 - P(X \le 4950) = 0.10564\ldots \approx 0.106$ **A1**

**(b)** invNorm$(0.10, 4800, 120)$ **(M1)**

$k = 4646.2\ldots \approx 4650$ bottles **A1**

*Note: Accept answers in range* $[4645, 4647]$. *Allow* $k = 4646$.

#### Q2 — Normal Distribution: Write Down and P(a < H < b) EASY

**(a)** Mean $= 72$ bpm; Standard deviation $= 8$ bpm **A1**

**(b)** $P(65 < H < 85)$ **M1**

$=$ normalcdf$(65, 85, 72, 8) = 0.76729\ldots \approx 0.767$ **A1**

*Note: FT from their (a) parameters.*

#### Q3 — Sample Mean Distribution: State Parameters EASY

**(a)** $\bar{M} \sim \text{N}\left(500,\ \frac{36}{36}\right) = \text{N}\left(500,\ \frac{6^2}{36}\right) = \text{N}(500,\ 1)$ **M1 A1**

Mean $= 500$ g; Variance $= 1$ g$^2$

**(b)** SE $= \frac{6}{\sqrt{36}} = \frac{6}{6} = 1$ g **A1**

*Note: SE = 1 follows directly from variance = 1; accept either route.*

#### Q4 — Linear Combination: Mean and Variance EASY

**(a)** $\text{E}(W) = \text{E}(X) + \text{E}(Y) = 30 + 20 = 50$ **A1**

**(b)** $\text{Var}(W) = \text{Var}(X) + \text{Var}(Y) = 4^2 + 3^2 = 16 + 9 = 25$ **A1**

**(c)** $W \sim \text{N}(50,\ 25)$ **A1**

*Note: Award A1 in (c) for correct distribution statement; accept* $\text{N}(50, 5^2)$.

### Q5 — Confidence Interval: Basic EASY

**(a)** 95% CI: $\bar{x} \pm 1.96 \times \dfrac{s_{n-1}}{\sqrt{n}} = 453 \pm 1.96 \times \dfrac{10}{\sqrt{25}}$ **M1**

$= 453 \pm 3.92 = (449.08, 456.92)$ **A1**

*Note: Accept* $(449, 457)$*; FT arithmetic errors.*

**(b)** The population of jar weights is normally distributed. **R1**

*Note: Accept "the sample is a random sample from a normal population." Do not accept "*$n$ *is large."*

### Q6 — Central Limit Theorem: Condition and Probability EASY

**(a)** The sample size $n = 64$ is sufficiently large ($n \geq 30$), so by the Central Limit Theorem $\bar{T}$ is approximately normally distributed. **R1**

**(b)** $\bar{T} \sim \mathrm{N}(18, 0.5^2)$

$P(\bar{T} > 19) = \text{normalcdf}(19, 10^{99}, 18, 0.5)$ **M1**

$= 0.02275\ldots \approx 0.0228$ **A1**

*Note:* $\sigma_{\bar{T}} = 4/\sqrt{64} = 0.5$ *may be derived or read from table.*

### Q7 — Inverse Normal and Interval Probability EASY

**(a)** $P(L > c) = 0.25 \Rightarrow P(L < c) = 0.75$ **M1**

$c = \text{invNorm}(0.75, 38, 5) = 41.37\ldots \approx 41.4$ cm **A1**

**(b)** $P(30 < L < 45) = \text{normalcdf}(30, 45, 38, 5) = 0.8804\ldots \approx 0.880$ **A1**

### Q8 — Confidence Interval: Read and Interpret EASY

**(a)** $\bar{x} = \dfrac{7240 + 8360}{2} = \dfrac{15600}{2} = 7800$ steps **A1**

**(b)** Margin of error $= \dfrac{8360 - 7240}{2} = \dfrac{1120}{2} = 560$ steps **A1**

**(c)** Since $7000 < 7240$, the value $7000$ lies outside the confidence interval; it is not plausible that the true mean is $7000$ steps. **R1**

### Q9 — Sample Mean Distribution and Probability EASY

**(a)** $\bar{C} \sim \mathrm{N}\left(7.2,\ \left(\dfrac{0.4}{\sqrt{16}}\right)^2\right) = \mathrm{N}(7.2,\ 0.1^2)$ **A1**

**(b)** $P(\bar{C} < 7.1) =$ normalcdf$(-10^{99}, 7.1, 7.2, 0.1)$ **M1**
$= 0.15866\ldots \approx 0.159$ **A1**

### Q10 — Combination of Two Independent Normals EASY

**(a)** E$(T) = 25 + 35 = 60$ min **A1**
Var$(T) = 4^2 + 5^2 = 16 + 25 = 41$
$\sigma_T = \sqrt{41} = 6.403\ldots \approx 6.40$ min **A1**
**(b)** $P(T > 70) =$ normalcdf$(70, 10^{99}, 60, \sqrt{41})$ **(M1)**
$= 0.05936\ldots \approx 0.0594$ **A1**

## Section B — Medium

### Q11 — Linear Combination: $2A + 3B$ MEDIUM

**(a)** E$(W) = 2(1000) + 3(500) = 2000 + 1500 = 3500$ g **A1**
**(b)** Var$(W) = 2^2(15^2) + 3^2(10^2) = 4(225) + 9(100) = 900 + 900 = 1800$ **M1**
$\sigma_W = \sqrt{1800} = 42.426\ldots \approx 42.4$ g **A1**
**(c)** $P(W > 3550) =$ normalcdf$(3550, 10^{99}, 3500, \sqrt{1800})$ **(M1)**
$= 0.11905\ldots \approx 0.119$ **A1**

### Q12 — Unbiased Estimates and 95% CI from Raw Data MEDIUM

**(a)** Using GDC: $\bar{x} = 45.5$ mm (exact: $45.5\overline{0}$ mm) **(M1) A1**
$s_{n-1} = 1.954\ldots$ mm; $s^2_{n-1} = 3.817\ldots \approx 3.82$ mm$^2$ **A1**
**(b)** 95% CI: $45.5 \pm 1.96 \times \dfrac{1.954\ldots}{\sqrt{12}}$ **M1**
$= 45.5 \pm 1.1047\ldots = (44.395\ldots,\ 46.604\ldots) \approx (44.4,\ 46.6)$ mm **A1**
*Note: Data: 43.2, 47.8, 45.1, 44.6, 46.3, 48.9, 42.7, 45.5, 47.2, 44.0, 46.8, 43.9. Sum* $= 546.0$; *mean* $= 45.5$. *FT from their mean and SD.*

### Q13 — Unknown $\mu$ and $\sigma$ from Two Probability Statements MEDIUM

**(a)** $P(M < 180) = 0.20 \Rightarrow \dfrac{180 - \mu}{\sigma} = -0.8416\ldots$ **A1**

$P(M > 260) = 0.15 \Rightarrow P(M < 260) = 0.85 \Rightarrow \dfrac{260 - \mu}{\sigma} = 1.0364\ldots$ **A1**

**(b)** Equations: $\mu - 0.8416\sigma = 180$ and $\mu + 1.0364\sigma = 260$ **M1**
Subtracting: $1.8780\sigma = 80 \Rightarrow \sigma = 42.6\ldots \approx 42.6$ g **A1**
$\mu = 180 + 0.8416 \times 42.6\ldots = 215.8\ldots \approx 216$ g **A1**
*Note: $z$-values:* invNorm$(0.20) = -0.8416$; invNorm$(0.85) = 1.0364$. *Accept solutions from GDC simultaneous solve. Final: $\mu \approx 216$ g, $\sigma \approx 42.6$ g.*

## Q14 — CLT Application with Non-Normal Population MEDIUM

**(a)** Since $n = 50 \geq 30$, by the Central Limit Theorem **R1**
$\bar{L} \sim \text{N}\left(85, \left(\frac{3}{\sqrt{50}}\right)^2\right) = \text{N}(85, 0.18)$ approximately. **A1**

**(b)** $\sigma_{\bar{L}} = 3/\sqrt{50} = 0.4243\ldots$
$P(84.2 < \bar{L} < 85.6) =$ normalcdf$(84.2, 85.6, 85, 3/\sqrt{50})$ **M1**
$= 0.8726\ldots \approx 0.873$ **A1**

## Q15 — Minimum Sample Size from Margin of Error MEDIUM

**(a)** Margin of error $= z^* \cdot \frac{\sigma}{\sqrt{n}} = 1.96 \times \frac{\sigma}{\sqrt{n}}$ **A1**

**(b)** Require $1.96 \times \frac{35}{\sqrt{n}} \leq 8$ **M1**
$\sqrt{n} \geq \frac{1.96 \times 35}{8} = \frac{68.6}{8} = 8.575$ **A1**
$n \geq 8.575^2 = 73.53\ldots$
Minimum $n = 74$ **A1**
*Note: Always round UP to the next integer for sample size questions.*

## Q16 — Difference of Two Independent Normals MEDIUM

**(a)** E$(D) = 178 - 165 = 13$ cm **A1**
Var$(D) = 7^2 + 6^2 = 49 + 36 = 85$
$\sigma_D = \sqrt{85} = 9.220\ldots \approx 9.22$ cm **A1**

**(b)** $P(D \geq 20) =$ normalcdf$(20, 10^{99}, 13, \sqrt{85})$ **M1**
$= 0.22311\ldots \approx 0.223$ **A1**
*Note: Variance of difference:* Var$(M - F) =$ Var$(M) +$ Var$(F)$ *(variances add for independent variables).*

### Q17 — Sample Mean Distribution: Graph, Probability, Inverse MEDIUM

**(a)** $\sigma_{\bar{L}} = \dfrac{25}{\sqrt{25}} = 5$ hours **A1**

**(b)** $P(\bar{L} > 425) =$ normalcdf$(425, 10^{99}, 420, 5)$ **(M1)**
$= 0.15866\ldots \approx 0.159$ **A1**

**(c)** $P(\bar{L} > k) = 0.05 \Rightarrow P(\bar{L} < k) = 0.95$ **M1**
$k =$ invNorm$(0.95, 420, 5) = 428.22\ldots \approx 428$ hours **A1**

### Q18 — CI from Summary Statistics with Plausibility Comment MEDIUM

**(a)** 95% CI: $497.4 \pm 1.96 \times \dfrac{8.2}{\sqrt{20}}$ **M1**
$= 497.4 \pm 3.592\ldots = (493.807\ldots,\ 500.992\ldots)$
$\approx (493.8,\ 501.0)$ g **A1**

**(b)** Since $500 \in (493.8,\ 501.0)$, the value $500$ g lies within the confidence interval. **R1**
It is plausible (at the 5% significance level) that the true mean net weight is $500$ g. **R1**
*Note: Both reasoning marks require a comparison AND a contextual conclusion. Accept "500 is inside the CI so the label claim is plausible."*

### Q19 — Weighted Combination $X + 2Y$ MEDIUM

**(a)** E$(S) = 24.0 + 2(24.0) = 72.0$ mm **A1**
Var$(S) = 0.3^2 + 2^2(0.5^2) = 0.09 + 4(0.25) = 0.09 + 1.00 = 1.09$
$\sigma_S = \sqrt{1.09} = 1.044\ldots \approx 1.04$ mm **A1**

**(b)** $P(S < 71) =$ normalcdf$(-10^{99}, 71, 72, \sqrt{1.09})$ **M1**
$= 0.16751\ldots \approx 0.168$ **A1**
*Note:* Var$(X + 2Y) =$ Var$(X) + 4\,$Var$(Y)$; *do not accept* Var$(X) + 2\,$Var$(Y)$.

### Q20 — Two Confidence Intervals from Table MEDIUM

**(a) Volta:** $18.4 \pm 1.96 \times \dfrac{2.1}{\sqrt{30}}$ **M1**
$= 18.4 \pm 0.7512\ldots = (17.649\ldots,\ 19.151\ldots) \approx (17.6,\ 19.2)$ hours **A1**

**(b) MaxPower:** $17.9 \pm 1.96 \times \dfrac{1.8}{\sqrt{30}}$ **M1**
$= 17.9 \pm 0.6438\ldots = (17.256\ldots,\ 18.544\ldots) \approx (17.3,\ 18.5)$ hours **A1**

*Note: Award M1 in each part for correct formula with their values. Accept slight rounding differences.*

## Section C — Hard

### Q21 — Show That: Distribution of $3P + 4Q$ HARD

**(a)** $\text{E}(T) = 3(12.0) + 4(8.0) = 36.0 + 32.0 = 68.0$ **A1**

$\text{Var}(T) = 3^2(0.8^2) + 4^2(0.5^2) = 9(0.64) + 16(0.25) = 5.76 + 4.00 = 9.76$ **M1 A1**

$\therefore T \sim \text{N}(68, 9.76)$ **AG**

**(b)** $P(T > m) = 0.10 \Rightarrow P(T < m) = 0.90$ **M1**

$m = \text{invNorm}(0.90, 68, \sqrt{9.76}) = \text{invNorm}(0.90, 68, 3.1241\ldots)$

$= 68 + 1.2816 \times 3.1241\ldots = 72.004\ldots \approx 72.0$ g **A1**

### Q22 — Unknown $\mu$ and $\sigma$; then Sample Mean Probability HARD

**(a)** $P(T < 45) = 0.30 \Rightarrow z_1 = \text{invNorm}(0.30) = -0.5244\ldots$

So $\dfrac{45 - \mu}{\sigma} = -0.5244 \Rightarrow \mu - 0.5244\sigma = 45$ **A1**

$P(T > 75) = 0.10 \Rightarrow z_2 = \text{invNorm}(0.90) = 1.2816\ldots$

So $\dfrac{75 - \mu}{\sigma} = 1.2816 \Rightarrow \mu + 1.2816\sigma = 75$ **A1**

Subtracting: $1.8060\sigma = 30 \Rightarrow \sigma = 16.61\ldots \approx 16.6$ min

$\mu = 45 + 0.5244 \times 16.61\ldots = 53.71\ldots \approx 53.7$ min **A1**

**(b)** $\bar{T} \sim \text{N}\left(53.71, \left(\dfrac{16.61}{\sqrt{16}}\right)^2\right) = \text{N}(53.71, 4.153^2)$ **(M1)**

$P(\bar{T} > 65) = \text{normalcdf}(65, 10^{99}, 53.71, 4.153) = 0.00324\ldots \approx 0.00324$ **A1**

*Note: FT from their $\mu$ and $\sigma$ in (a). Accept $P(\bar{T} > 65) \approx 0.00324$ or $3.24 \times 10^{-3}$.*

### Q23 — 99% CI Sample Size then CI Calculation HARD

**(a)** Total width $\leq 60 \Rightarrow$ margin of error $\leq 30$ **M1**

$2.576 \times \dfrac{180}{\sqrt{n}} \leq 30$ **A1**

$\sqrt{n} \geq \dfrac{2.576 \times 180}{30} = \dfrac{463.68}{30} = 15.456$

$n \geq 238.89\ldots \Rightarrow n_{\text{min}} = 239$ **A1**

**(b)** 99% CI with $n = 250$, $\bar{x} = 1045$, $\sigma = 180$:

$1045 \pm 2.576 \times \dfrac{180}{\sqrt{250}} = 1045 \pm 2.576 \times 11.384\ldots = 1045 \pm 29.325\ldots$ **M1**

$= (1015.674\ldots,\ 1074.325\ldots) \approx (1016,\ 1074)$ mg **A1**

*Note: Accept* $(1015.7, 1074.3)$ *or equivalent 4 sf bounds.*

## Q24 — Weighted Combination with Nutritional Standard HARD

**(a)** $\mathrm{E}(W) = 0.6(22) + 0.4(15) = 13.2 + 6.0 = 19.2$ g **A1**

$\mathrm{Var}(W) = 0.6^2(2^2) + 0.4^2(1.5^2) = 0.36(4) + 0.16(2.25) = 1.44 + 0.36 = 1.80$ **M1 A1**

**(b)** $W \sim \mathrm{N}(19.2, 1.80)$; $\sigma_W = \sqrt{1.80} = 1.3416\ldots$

$P(W > 19) = \text{normalcdf}(19, 10^{99}, 19.2, 1.3416)$ **M1**

$= 0.5592\ldots \approx 0.559$ **A1**

## Q25 — CI Width and Minimum Sample Size HARD

**(a)** 95% CI with $n = 16$, $\bar{x} = 14.3$, $s_{n-1} = 2.4$:

$14.3 \pm 1.96 \times \dfrac{2.4}{\sqrt{16}} = 14.3 \pm 1.96 \times 0.6 = 14.3 \pm 1.176$ **M1**

$= (13.124,\ 15.476) \approx (13.1,\ 15.5)$ cm **A1**

**(b)** Width $< 1.5 \Rightarrow 2 \times 1.96 \times \dfrac{2.4}{\sqrt{n}} < 1.5$ **M1**

$\sqrt{n} > \dfrac{2 \times 1.96 \times 2.4}{1.5} = \dfrac{9.408}{1.5} = 6.272$

$n > 39.33\ldots$ **A1**

Minimum $n = 40$ **A1**

*Note: Must round UP. Do not accept* $n = 39$.

## Q26 — CLT Reject Probability and Inverse Mean HARD

**(a)** Since $n = 40 \geq 30$, by the Central Limit Theorem **R1**

$\bar{M} \sim \mathrm{N}\left(185, \left(\dfrac{20}{\sqrt{40}}\right)^2\right) = \mathrm{N}(185, 10)$ approximately **A1**

**(b)** $P(\bar{M} < 180) = \text{normalcdf}(-10^{99}, 180, 185, \sqrt{10})$ **M1**

$= 0.05691\ldots \approx 0.0569$ **A1**

**(c)** Require $P(\bar{M} < 180) = 0.01$ **M1**

$\dfrac{180 - \mu_0}{\sqrt{10}} = \text{invNorm}(0.01) = -2.3263\ldots$

$\mu_0 = 180 + 2.3263 \times \sqrt{10} = 180 + 7.357\ldots = 187.36\ldots \approx 187$ g **A1**

### Q27 — Unknown Parameter $\mu$ from Probability Condition HARD

**(a)** $\text{E}(D) = \mu - 10$ **A1**

**(b)** $\text{Var}(D) = \text{Var}(X) + \text{Var}(Y) = 5^2 + 3^2 = 25 + 9 = 34$ **M1**

*Note:* $\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y)$; signs do not affect variance.

$\sigma_D = \sqrt{34}$ **AG**

*Note: Fully justified step required for AG.*

**(c)** $P(D > 8) = 0.20 \Rightarrow P(D < 8) = 0.80$ **M1**

$\dfrac{8 - (\mu - 10)}{\sqrt{34}} = \text{invNorm}(0.80) = 0.8416\ldots$

$18 - \mu = 0.8416 \times \sqrt{34} = 4.908\ldots$

$\mu = 18 - 4.908\ldots = 13.09\ldots \approx 13.1$ **A1**

### Q28 — Two CIs from Table with Comparison HARD

**(a) ArcoTyre:** $62.4 \pm 1.96 \times \dfrac{4.8}{\sqrt{18}}$ **M1**

$= 62.4 \pm 2.217\ldots = (60.182\ldots,\ 64.617\ldots) \approx (60.2,\ 64.6)$ 000 km **A1**

**(b) RoadKing:** $58.9 \pm 1.96 \times \dfrac{5.6}{\sqrt{22}}$ **M1**

$= 58.9 \pm 2.341\ldots = (56.558\ldots,\ 61.241\ldots) \approx (56.6,\ 61.2)$ 000 km **A1**

**(c)** The two confidence intervals overlap (ArcoTyre's lower bound $\approx 60.2$ is above RoadKing's upper bound $\approx 61.2$ — wait, they share region $[60.2, 61.2]$). Since the intervals overlap, the data do not provide conclusive evidence that the two brands have different mean tread wear. **R1**

*Note: Award R1 for noting the intervals overlap AND stating the implication for the comparison.*

### Q29 — Maximum $\sigma$ from Sample Mean Condition HARD

**(a)** $\bar{V} \sim \text{N}\left(200, \dfrac{\sigma^2}{25}\right)$; standard deviation $= \dfrac{\sigma}{5}$ **A1**

**(b)** $P(\bar{V} > 202) \le 0.05$

$\dfrac{202 - 200}{\sigma/5} \ge 1.645$ **M1**

$\dfrac{10}{\sigma} \ge 1.645 \Rightarrow \sigma \le \dfrac{10}{1.645}$ **AG**

**(c)** $\sigma_{\text{max}} = \dfrac{10}{1.645} = 6.079\ldots \approx 6.08$ ml **A1**

**(d)** $\sigma = 6.0$; $\sigma_{\bar{V}} = 6.0/5 = 1.2$

$P(198 < \bar{V} < 203) =$ normalcdf$(198, 203, 200, 1.2)$ **(M1)**

$= 0.8771\ldots \approx 0.877$ **A1**

**Q30 — Full Synthesis: Combination, Sample Mean, CLT** HARD

**(a)** $\text{E}(C) = 480 + 720 = 1200$ kcal **A1**

$\text{Var}(C) = 60^2 + 90^2 = 3600 + 8100 = 11700$

$\sigma_C = \sqrt{11700} = 108.167\ldots \approx 108$ kcal **A1**

**(b)** $\bar{C} \sim \text{N}\left(1200,\ \dfrac{11700}{36}\right) = \text{N}(1200,\ 325)$; $\sigma_{\bar{C}} = \sqrt{325} = 18.028\ldots$

$P(1170 < \bar{C} < 1230) =$ normalcdf$(1170, 1230, 1200, \sqrt{325})$ **M1**

$= 0.9035\ldots \approx 0.903$ (or $0.904$) **A1**

**(c)** The Central Limit Theorem allows $\bar{C}$ to be treated as approximately normal; the condition is that the sample size $n = 36$ is sufficiently large ($n \geq 30$). **R1**