Counting and statistics - Consolidation test 5

10 questions | 54 marks | 75 minutes

Connect counting, sampling and data summaries to a quality-control decision. Explain model assumptions and support comparisons numerically.

Equivalent correct methods earn the corresponding marks.

1. An access code is either two distinct letters from A-F followed by one digit 0-9, or one letter from A-F followed by three distinct digits 0-9. A code may begin with any listed letter. [5 marks]

(a) Count the codes in each format. [3 marks]

First format: $6 \times 5 \times 10 = 300$.

Second format: $6 \times 10 \times 9 \times 8 = 4320$.

Answer: 300 and 4,320.

(b) Count all codes and explain why adding is valid. [2 marks]

The formats have different lengths, so no code is counted in both.

$300 + 4320 = 4620$.

Answer: 4,620.

2. A workshop has 6 trainees and 4 mentors. A working group has 4 members. [6 marks]

(a) Count groups with at least one mentor. [3 marks]

All groups $\binom{10}{4} = 210$.

All-trainee groups $\binom{6}{4} = 15$.

Subtract $210 - 15 = 195$.

Answer: 195.

(b) Count groups with exactly two mentors and two trainees, together with a choice of one group leader. [3 marks]

Choose members in $\binom{4}{2}\binom{6}{2} = 6 \times 15 = 90$ ways.

Any of the 4 members can lead, so $90 \times 4 = 360$.

Answer: 360.

3. A model uses $(2x+3)^4$. [5 marks]

(a) Find the coefficient of x^2. [3 marks]

$\binom{4}{2} \times 2^2 \times 3^2 = 6 \times 4 \times 9 = 216.$

Answer: 216.

(b) Find the sum of its coefficients without a full expansion. [2 marks]

At $x = 1$, the value is $(2+3)^4 = 625$.

Answer: 625.

4. A school compares bus and tram journey times. A volunteer survey records times from 12 bus users and 80 tram users. [6 marks]

(a) Give two reasons to be cautious about saying one transport mode is better for every student. [3 marks]

The samples may not represent all users.

An observed mode association is not proof that the mode causes all differences.

Answer: Voluntary response may be biased; different routes, distances or departure times may explain differences.

(b) Describe a more defensible study and name the numerical variable. [3 marks]

Define the population and sampling frame before selection.

Record durations in a common unit and avoid choosing only convenient users.

Answer: Use a planned random sample within each mode, record comparable route/distance/time information and report limitations; journey duration is continuous numerical.

5. The table gives weekly exercise-session counts. [6 marks]

Sessions	0	1	2	3	4
Frequency	2	4	6	4	2

(a) Find the number of observations, mean, median and mode. [3 marks]

Total frequency 18; weighted total 36.

The ninth and tenth observations are 2;2 has the largest frequency 6.

Answer: $n = 18$; mean 2; median 2; mode 2.

(b) Using the sample formula, calculate the standard deviation to 3 decimal places. [3 marks]

About mean 2, the weighted squared-deviation sum is $2(4) + 4(1) + 6(0) + 4(1) + 2(4) = 24$.

$s = \sqrt{\frac{24}{17}} = 1.188177\ldots$

Answer: $s \approx 1.188$.

6. One sample of 10 journeys has mean 40 min. Another sample of 15 journeys has mean 50 min. [5 marks]

(a) Find the combined mean. [3 marks]

Combined total $10 \times 40 + 15 \times 50 = 1150$.

Divide by 25:46.

Answer: 46 min.

(b) Can the combined median be found from these sample sizes and means alone? Explain. [2 marks]

Means and sample sizes give totals, not the ordered individual observations needed for a median.

Answer: No.

7. Ordered measurements are 12,14,15,16,17,18,20,40. Use median-of-halves quartiles. [6 marks]

(a) Find median, IQR and range. [3 marks]

Median $= \frac{16+17}{2} = 16.5$.

$Q_1 = \frac{14+15}{2} = 14.5$; $Q_3 = \frac{18+20}{2} = 19$.

IQR 4.5; range $40 - 12 = 28$.

Answer: Median 16.5; IQR 4.5; range 28.

(b) The 40 was a recording error and should be 22. Which of these three summaries change? [3 marks]

The middle values and half medians stay the same.

The new maximum is 22 and $22 - 12 = 10$.

Answer: Only range changes, to 10.

8. A sample of 7 packages has mean 500 g and sum of squared deviations 54 g². [4 marks]

(a) Find the sample variance and standard deviation. [2 marks]

$\frac{54}{7-1} = 9$; $\sqrt{9} = 3$.

Answer: Variance 9 g²; $s = 3$ g.

(b) Each package has a constant 20 g container. Find the mean and s of the contents. [2 marks]

Subtracting the same constant shifts centre but preserves spread.

Answer: Mean 480 g; $s = 3$ g.

9. A process is normally modelled with mean 20 mm and sd 0.5 mm. Items are within tolerance from 19.25 to 20.75 mm. Supplied $\Phi(1.5) = 0.9331928$. [5 marks]

(a) Find the modelled proportion within tolerance. [3 marks]

Both cutoffs are 1.5 standard deviations from the mean.

By symmetry the central area is $2\Phi(1.5) - 1$.

Answer: 0.8663856, about 86.64%.

(b) Estimate the expected count outside tolerance among 2,000 items. [2 marks]

$2000 \times (1 - 0.8663856) = 267.2288$.

Answer: 267.2288, approximately 267.

10. A proposed process has normal mean 20 mm and sd 0.3 mm. The tolerance is 19.25 to 20.75 mm. A previous normal model with the same mean and sd 0.5 mm predicts 0.8663856 within tolerance. Supplied $\Phi(2.5) = 0.9937903$. [6 marks]

(a) Find its predicted within-tolerance proportion and compare with the preceding model. [3 marks]

The cutoffs are ± 2.5 standard deviations.

$2 \times 0.9937903 - 1 = 0.9875806$.

$0.9875806 - 0.8663856 = 0.1211950$.

Answer: 0.9875806, about 98.76%; an increase of about 12.12 percentage points.

(b) Explain what would need checking before promising that proportion in production. [3 marks]

A model prediction depends on valid assumptions.

An expected long-run proportion does not guarantee any particular batch result.

Answer: Check the process data, stable mean/spread, normal-shape assumption and representative sampling; actual batches vary.