Past-paper walk-through

The Normal Distribution ▪ 1.10

eXercise

Questions in this set

- 2026 Q3
- 2023 Q6
- 2022 Q3
- 2019 Q5
- 2017 Q3
- 2014 Q3

Reading the mark scheme

- **B1** a mark that stands on its own – the point is either made or not.
- **M1** a *method* mark: the right approach, even if the number is wrong.
- **A1** an *answer* mark, and it usually needs its M mark first.
- **C1** a working mark on the way to the answer.
- // separates answers the examiner accepts equally.
- **ecf** = error carried forward: a later part is marked on your own earlier answer.
- **owtte** = "or words to that effect" – the wording need not match.

Question 3

2026 ▪ Q3

A certain sports team has its team song performed by a different musician at the beginning of each of its games. The time it takes for the team song to be performed by different musicians can be modeled using a normal distribution with mean $\mu = 109$ seconds and standard deviation $\sigma = 16$ seconds. All performances of the team song are independent.

(a) What is the probability that a randomly selected performance of the team song will take longer than 120 seconds? Show your work.

(b) Ten performances of the team song will be randomly selected. Let the random variable X represent the number of games at which the performance of the team song takes longer than 120 seconds. Find $P(X \geq 3)$. Show your work.

Question 3

(c)(i) Ben will attend all of the games for this sports team. Let the random variable Y represent the number of games Ben will attend until a performance of the team song takes longer than 120 seconds.
Calculate the mean of Y. Show your work.
(c)(ii) Calculate the standard deviation of Y. Show your work.
(d) Interpret the standard deviation calculated in part C (ii) in context.

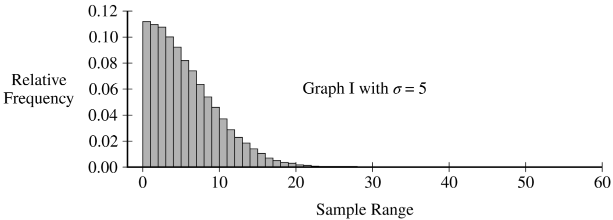
Question 6

2023 ▪ Q6

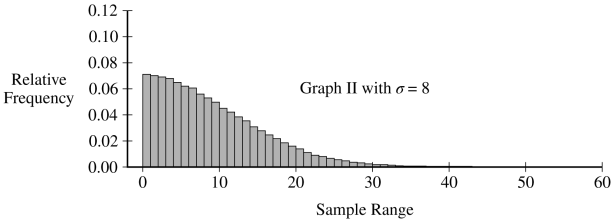
A jewelry company uses a machine to apply a coating of gold on a certain style of necklace. The amount of gold applied to a necklace is approximately normally distributed. When the machine is working properly, the amount of gold applied to a necklace has a mean of 300 milligrams (mg) and standard deviation of 5 mg.

(a) A necklace is randomly selected from the necklaces produced by the machine. Assuming that the machine is working properly, calculate the probability that the amount of gold applied to the necklace is between 296 mg and 304 mg.

Question 6



Relative Frequency
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
10
20
30
40
50
60
Sample Range
Graph I with σ = 5

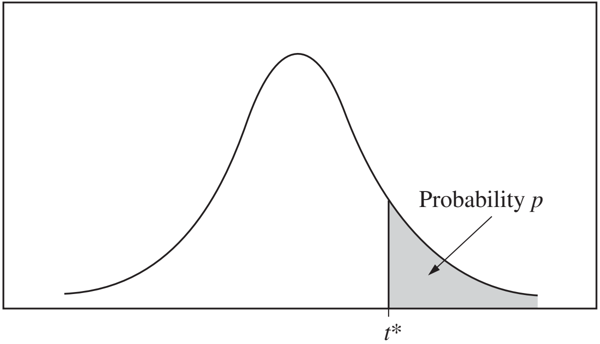


Relative Frequency
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
10
20
30
40
50
60
Sample Range
Graph II with σ = 8

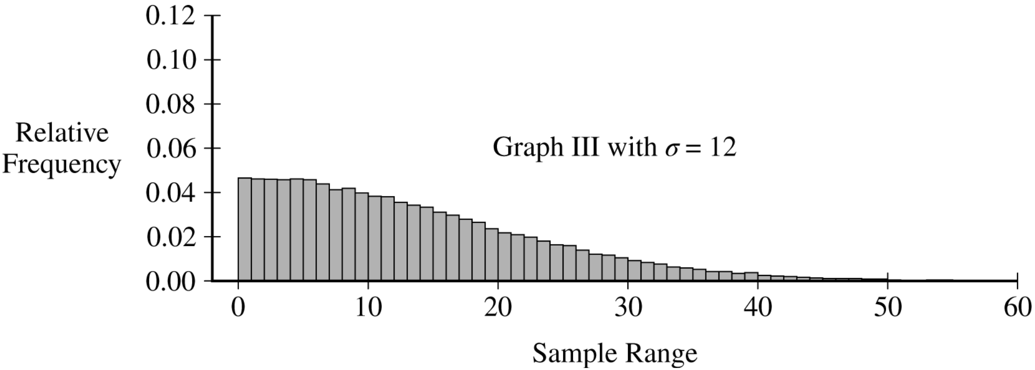
Question 6

3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Table entry for p and C is the point t^* with probability p lying above it and probability C lying between $-t^*$ and t^*.



Question 6



Relative Frequency
0.12
0.10
0.08
0.06
0.04
0.02
0.00
Graph III with σ = 12
0
10
20
30
40
50
60
Sample Range

Question 6

2023 ▪ Q6

A jewelry company uses a machine to apply a coating of gold on a certain style of necklace. The amount of gold applied to a necklace is approximately normally distributed. When the machine is working properly, the amount of gold applied to a necklace has a mean of 300 milligrams (mg) and standard deviation of 5 mg.

(b)(i) The jewelry company wants to make sure the machine is working properly. Each day, Cleo, a statistician at the jewelry company, will take a random sample of the necklaces produced that day. Each selected necklace will be melted down and the amount of the gold applied to that necklace will be determined. Because a necklace must be destroyed to determine the amount of gold that was applied, Cleo will use random samples of size $n = 2$ necklaces.

Question 6

Cleo starts by considering the mean amount of gold being applied to the necklaces. After Cleo takes a random sample of $n = 2$ necklaces, she computes the sample mean amount of gold applied to the two necklaces.
Suppose the machine is working properly with a population mean amount of gold being applied of 300 mg and a population standard deviation of 5 mg.
Calculate the probability that the sample mean amount of gold applied to a random sample of $n = 2$ necklaces will be greater than 303 mg.
(b)(ii) Suppose Cleo took a random sample of $n = 2$ necklaces that resulted in a sample mean amount of gold applied of 303 mg. Would that result indicate that the population mean amount of gold being applied by the machine is different from 300 mg? Justify your answer without performing an inference procedure.

Question 6

2023 ▪ Q6

A jewelry company uses a machine to apply a coating of gold on a certain style of necklace. The amount of gold applied to a necklace is approximately normally distributed. When the machine is working properly, the amount of gold applied to a necklace has a mean of 300 milligrams (mg) and standard deviation of 5 mg.

(c)(i) Now, Cleo will consider the variation in the amount of gold the machine applies to the necklaces. Because of the small sample size, $n = 2$, Cleo will use the sample range of the data for the two randomly selected necklaces, rather than the sample standard deviation.

Cleo will investigate the behavior of the range for samples of size $n = 2$. She will simulate the sampling distribution of the range of the amount of gold applied to two randomly sampled necklaces. Cleo generates $100{,}000$ random samples of size $n = 2$ independent values from a normal distribution with mean $\mu = 300$ and standard

Question 6

deviation $\sigma = 5$. The range is calculated for the two observations in each sample. The simulated sampling distribution of the range is shown in Graph I. This process is repeated using $\sigma = 8$, as shown in Graph II, and again using $\sigma = 12$, as shown in Graph III.

[Graph I with $\sigma = 5$: histogram titled "Graph I with $\sigma = 5$"; x-axis "Sample Range" from 0 to 60; y-axis "Relative Frequency" from 0.00 to 0.12. Distribution is strongly right-skewed, peak (about 0.11-0.12) near sample range 0-2, decreasing steadily to near 0 by about range 25-30.]

[Graph II with $\sigma = 8$: histogram titled "Graph II with $\sigma = 8$"; x-axis "Sample Range" from 0 to 60; y-axis "Relative Frequency" from 0.00 to 0.12. Distribution is right-skewed, peak (about 0.07) near sample range 0-2, decreasing steadily to near 0 by about range 40-45.]

Question 6

[Graph III with $\sigma = 12$: histogram titled "Graph III with $\sigma = 12$"; x-axis "Sample Range" from 0 to 60; y-axis "Relative Frequency" from 0.00 to 0.12. Distribution is right-skewed, peak (about 0.047) near sample range 0-2, decreasing gradually, with a longer tail extending out to about range 55-60.]

Use the information in the graphs to complete the following.

Describe the sampling distribution of the sample range for random samples of size $n = 2$ from a normal distribution with standard deviation $\sigma = 5$, as shown in Graph I.

(c)(ii) Describe how the sampling distribution of the sample range for samples of size $n = 2$ changes as the value of the population standard deviation σ increases.

Question 6

2023 ▪ Q6

A jewelry company uses a machine to apply a coating of gold on a certain style of necklace. The amount of gold applied to a necklace is approximately normally distributed. When the machine is working properly, the amount of gold applied to a necklace has a mean of 300 milligrams (mg) and standard deviation of 5 mg.

(d)(i) Recall that Cleo needs to consider both the mean and standard deviation of the amount of gold applied to necklaces to determine whether the machine is working properly. Suppose that one month later, Cleo is again checking the machine to make sure it is working properly. Cleo takes a random sample of 2 necklaces and calculates the sample mean amount of gold applied as 303 mg and the sample range as 10 mg. Recall that the machine is working properly if the amount of gold applied to the necklaces has a mean of 300 mg and standard deviation of 5 mg.

Question 6

Consider Cleo's range of 10 mg from the sample of size $n = 2$. If the machine is working properly with a standard deviation of 5 mg, is a sample range of 10 mg unusual? Justify your answer.

(d)(ii) Do Cleo's sample mean of 303 mg and range of 10 mg indicate that the machine is not working properly? Explain your answer.

Solution 6

(a) Mark scheme

- Let X represent the amount of gold (mg) applied to a randomly selected necklace. Assuming the machine works properly, X has an approximately normal distribution with mean 300 mg and standard deviation 5 mg. $P(296 < X < 304) = P(X<304) - P(X\leq 296) = P(Z < (304-300)/5) - P(Z \leq (296-300)/5) = P(Z<0.8) - P(Z\leq -0.8) \approx 0.7881 - 0.2119 \approx 0.5763$. Full credit needs all 3 components: (1) indicates use of a normal distribution with mean 300 and standard deviation 5 (may be shown graphically, via calculator normalcdf syntax, in words, in standard notation N(300,5), or via a labeled z-score calculation), (2) specifies the correct event (boundary values 296 and 304, correct direction) or an event consistent with component 1, (3) provides the correct probability 0.5763, or a probability

Solution 6

consistent with components 1 and 2. Partial credit if only 2 of the 3 components are satisfied. Minor errors in statistical notation may be ignored (though considered poor communication under holistic scoring). If the only error is a reversed z-score numerator (300—296 or 300—304), the response is scored partially correct.

Solution 6

(b)(i) Mark scheme

- If the machine is working properly, the sample mean amount of gold $\bar{X}$ (for a sample of n=2 necklaces) has a sampling distribution that is approximately normal with mean 300 mg and standard deviation $5/\sqrt{2} \approx 3.5355$ mg. Then $P(\bar{X} > 303) = P(Z > (303-300)/3.5355) \approx P(Z > 0.8485) \approx 0.198$. NOTE: parts (b)(i) and (b)(ii) are scored TOGETHER as ONE holistic unit in the official scoring guidelines, so this leaf carries this part's full mark. Full credit for the combined part (b) needs components 1, 4, and 5 AND at least one of components 2 or 3: (1) in part (b)(i), indicates use of a normal/approximately-normal distribution and identifies the correct parameter values (mean 300, SD $5/\sqrt{2} \approx 3.5355$) with work shown or a correct formula, (2) in part (b)(i), specifies the correct event (boundary

Solution 6

value and direction: $\bar{X}>303$), (3) in part (b)(i), provides the correct probability 0.198 (or one consistent with components 1-2), (4) in part (b)(ii), indicates whether observing a sample mean of 303 mg would provide convincing evidence the population mean is not 300 mg, consistent with the response to (b)(i) [graded together with leaf 6(b)(ii)], (5) in part (b)(ii), provides justification based on the probability obtained in (b)(i) or a correctly-computed probability in (b)(ii). An alternate solution using the TOTAL amount of gold on the 2 necklaces (N(600, 7.071), P(total>606)) earning full component-1/2/3 credit is also acceptable.

Solution 6

(b)(ii) Mark scheme

- Observing a sample mean amount of 303 mg would NOT provide convincing evidence that the population mean amount of gold being applied by the machine is something other than 300 mg, because the probability of observing a sample mean that differs from 300 mg by 3 mg or more (in either direction) is large - around $0.198(2) = 0.396$ - so a difference this size is not statistically surprising if the machine were working properly. This leaf's credit is bundled with 6(b)(i) in the official scoring guidelines (Part (b) of the SG scores both sub-parts together as one holistic unit), so it carries 0 of this question's marks on its own - but a correct, well-justified 'no' answer, consistent with the probability found in (b)(i), is still expected to complete credit for part (b).

Solution 6

(c)(i) Mark scheme

- Cleo simulates the sampling distribution of the sample range (max−min) for random samples of size n=2 from a normal distribution with standard deviation σ=5 mg (100,000 simulated samples). The resulting sampling distribution of the sample range is skewed to the right. Almost all simulated values of the range fall between 0 mg and about 25-30 mg, and the center of the distribution is about 6 mg. NOTE: parts (c)(i) and (c)(ii) are scored TOGETHER as ONE holistic unit in the official scoring guidelines, so this leaf carries this part's full mark. Full credit for the combined part (c) needs 4 or 5 of the following 5 components: (1) in part (c)(i), describes the shape as positively skewed (skewed right), (2) in part (c)(i), describes the center of the distribution as about 6 mg (any value between 4

Solution 6

and 10 mg is a reasonable center description), (3) in part (c)(i), describes the spread as having most values between 0 mg and 30 mg (or IQR between 5 and 7 mg inclusive), (4) in part (c)(ii), indicates the mean/center of the sample-range distribution increases as the population SD σ increases [graded together with leaf 6(c)(ii)], (5) in part (c)(ii), indicates the variation/spread of the sample-range distribution also increases as σ increases. Partial credit if only 3 of the 5 components are satisfied.

Solution 6

(c)(ii) Mark scheme

- As the value of the population standard deviation σ increases, the variation (spread) in the distribution of the sample range increases, and the mean (center) of the distribution of the sample range also increases. This leaf's credit is bundled with 6(c)(i) in the official scoring guidelines (Part (c) of the SG scores both sub-parts together as one holistic unit), so it carries 0 of this question's marks on its own - but stating BOTH that the center increases AND that the spread increases (as σ increases) is still expected to complete credit for part (c). Comments on skewness are not required here and should be treated as extraneous.

Solution 6

(d)(i) Mark scheme

- No, a sample range of 10 mg is not unusual if the machine is working properly with a standard deviation of 5 mg. Although observing a sample range around 10 mg or greater is much more likely if the population SD is 8 mg or 12 mg, the sampling distribution of the sample range (n=2) for a normal distribution with σ=5 mg shows that 10 mg is not an unusual value - there is about a 20% chance that a random sample of 2 necklaces would yield a range of 10 mg or more when the machine is working properly. NOTE: parts (d)(i) and (d)(ii) are scored TOGETHER as ONE holistic unit in the official scoring guidelines, so this leaf carries this part's full mark. Full credit for the combined part (d) needs 4 or 5 of the following 5 components: (1) in part (d)(i), indicates a sample range of 10 mg

Solution 6

is not unusual, (2) in part (d)(i), justifies this by indicating the $\sigma=5$ sampling-distribution graph shows values of 10 mg or greater occur fairly often, (3) in part (d)(ii), indicates there is not convincing evidence the machine is not working properly [graded together with leaf 6(d)(ii)], (4) in part (d)(ii), justifies this by indicating a sample mean of 303 mg would not be unusual (e.g. it is less than one SD from 300, or consistent with the $\approx$0.198 probability from part (b)), (5) in part (d)(ii), justifies this by indicating a sample range of 10 mg is not unusual, consistent with (d)(i). Partial credit if at least 3 of the 5 components are satisfied.

Solution 6

(d)(ii) Mark scheme

- No, Cleo's sample mean of 303 mg and range of 10 mg do NOT indicate that the machine is not working properly. As noted in part (b)(i), the probability that the sample mean would be equal to or greater than 303 mg when the machine is working properly is almost 20%, so a sample mean of 303 mg is not unusual (it is also less than one standard deviation, $5/\sqrt{2}\approx 3.5355$ mg, away from 300 mg). As indicated in part (d)(i), the probability of a range of 10 mg or greater when the population SD is 5 mg is also about 20%, so it too is not unusual. There is not statistically significant evidence to show the machine is not working properly. This leaf's credit is bundled with 6(d)(i) in the official scoring guidelines (Part (d) of the SG scores both sub-parts together as one holistic unit), so it carries 0 of this

Solution 6

question's marks on its own - but this correct 'no' conclusion, justified by BOTH the mean's and the range's probabilities not being unusual (consistent with (d)(i)), is still expected to complete credit for part (d). Context (amount of gold in necklaces) is not required in parts (a)-(d) but is considered good communication.

Question 3

2022 ▪ Q3

A machine at a manufacturing company is programmed to fill shampoo bottles such that the amount of shampoo in each bottle is normally distributed with mean 0.60 liter and standard deviation 0.04 liter. Let the random variable A represent the amount of shampoo, in liters, that is inserted into a bottle by the filling machine.
(a) A bottle is considered to be underfilled if it has less than 0.50 liter of shampoo. Determine the probability that a randomly selected bottle of shampoo will be underfilled. Show your work.

Question 3

2022 ▪ Q3

A machine at a manufacturing company is programmed to fill shampoo bottles such that the amount of shampoo in each bottle is normally distributed with mean 0.60 liter and standard deviation 0.04 liter. Let the random variable A represent the amount of shampoo, in liters, that is inserted into a bottle by the filling machine.

(b)(i) After the bottles are filled, they are placed in boxes of 10 bottles per box. After the bottles are placed in the boxes, several boxes are placed in a crate for shipping to a beauty supply warehouse. The manufacturing company's contract with the beauty supply warehouse states that one box will be randomly selected from a crate. If 2 or more bottles in the selected box are underfilled, the entire crate will be rejected and sent back to the manufacturing company.

Question 3

The beauty supply warehouse manager is interested in the probability that a crate shipped to the warehouse will be rejected. Assume that the amounts of shampoo in the bottles are independent of each other.
Define the random variable of interest for the warehouse manager and state how the random variable is distributed.
(b)(ii) Determine the probability that a crate will be rejected by the warehouse manager. Show your work.

Question 3

2022 ▪ Q3

A machine at a manufacturing company is programmed to fill shampoo bottles such that the amount of shampoo in each bottle is normally distributed with mean 0.60 liter and standard deviation 0.04 liter. Let the random variable A represent the amount of shampoo, in liters, that is inserted into a bottle by the filling machine.

(c) To reduce the number of crates rejected by the beauty supply warehouse manager, the manufacturing company is considering adjusting the programming of the filling machine so that the amount of shampoo in each bottle is normally distributed with mean 0.56 liter and standard deviation 0.03 liter.

Would you recommend that the manufacturing company use the original programming of the filling machine or the adjusted programming of the filling machine? Provide a statistical justification for your choice.

Solution 3

(a) Mark scheme

- Let A be the amount of shampoo (in liters) in a randomly selected bottle; A follows a normal distribution with mean 0.6 liter and standard deviation 0.04 liter. Then $P(A < 0.5) = P\left(Z < \frac{0.5 - 0.6}{0.04} = -2.5\right) \approx 0.0062$. Full credit needs all three components: (1) indicates use of a normal (or approximately normal) distribution with the correct parameter values (mean 0.6, standard deviation 0.04), (2) specifies the correct event (boundary value 0.5 and the correct direction, 'less than') consistent with component 1, (3) provides the correct probability of 0.0062 (or a probability consistent with components 1 and 2). An arithmetic/transcription error can be ignored if correct work is shown; reversing

Solution 3

the *z*-score numerator $(0.6 - 0.5$ instead of $0.5 - 0.6)$ as the ONLY error earns partial credit.

Solution 3

(b)(i)

Mark scheme

- The random variable of interest, X, is the number of underfilled bottles in a box of 10 bottles. The distribution of X is binomial with parameters $n = 10$ and $p = 0.0062$ (the probability from part (a)). This is the first of four components shared with part (b-ii) as a single scored part: full credit for the shared point additionally requires (from (b-ii)) supporting work identifying the event of interest ('at least two underfilled bottles') and the correct final probability of approximately 0.0017.

Solution 3

(b)(ii) Mark scheme

- A box's crate is rejected if 2 or more of its 10 bottles are underfilled, i.e. $X \geq 2$. $P(X \geq 2) = 1 - P(X \leq 1) =$ $1 - \binom{10}{1}(0.0062)^1(0.9938)^9 - \binom{10}{0}(0.0062)^0(0.9938)^{10} \approx 0.0017$. This is shared with part (b-i) as a single scored part covering four components: (1) part (b-i) defines the random variable as the number of underfilled bottles in a box of 10, (2) indicates the binomial distribution with $n = 10$ and $p = 0.0062$ (may appear in either (b-i) or (b-ii)), (3) provides supporting work for this part-(b-ii) calculation that identifies the event of interest (at least two underfilled), (4) calculates the correct probability of approximately 0.0017 (or a probability consistent with the

Solution 3

response to (a) or (b-i)). An arithmetic/transcription error can be ignored if correct work is shown.

Solution 3

(c) Mark scheme

- The company should continue using the ORIGINAL programming of the filling machine. Original programming: $P(A < 0.5) = P\left(Z < \dfrac{0.5 - 0.60}{0.04}\right) = P(Z < -2.5) \approx 0.0062$. Adjusted programming (mean 0.56 L, sd 0.03 L): $P(A < 0.5) = P\left(Z < \dfrac{0.5 - 0.56}{0.03}\right) = P(Z < -2.0) \approx 0.02275$. Because the probability of an underfilled bottle is greater under the adjusted programming (0.023 vs. 0.0062), it would result in more rejected shipments (using the binomial model from part (b), the crate-rejection probability rises from about 0.0017 under

Solution 3

the original programming to about 0.0206 under the adjusted programming), so the company should keep the original programming. Full credit needs both: (1) correctly calculates the probability (or *z*-score) of underfilling for the adjusted programming, (2) provides a correct conclusion about which programming to recommend, based on a comparison of the original and adjusted probabilities (or *z*-scores) – a bare 'yes'/'no' without a stated recommendation does not satisfy component 2. A response that correctly uses the binomial distribution to compute and compare crate-rejection probabilities (0.0017 vs. 0.0206) with a correct conclusion also earns full credit.

Question 5

2019 ▪ Q5

A company that manufactures smartphones developed a new battery that has a longer life span than that of a traditional battery. From the date of purchase of a smartphone, the distribution of the life span of the new battery is approximately normal with mean 30 months and standard deviation 8 months. For the price of $50, the company offers a two-year warranty on the new battery for customers who purchase a smartphone. The warranty guarantees that the smartphone will be replaced at no cost to the customer if the battery no longer works within 24 months from the date of purchase.

(a) In how many months from the date of purchase is it expected that 25 percent of the batteries will no longer work? Justify your answer.

(b) Suppose one customer who purchases the warranty is selected at random. What is the probability that the customer selected will require a replacement within 24 months from the date of purchase because the battery no longer works?

Question 5

(c) The company has a gain of \$50 for each customer who purchases a warranty but does not require a replacement. The company has a loss (negative gain) of \$150 for each customer who purchases a warranty and does require a replacement. What is the expected value of the gain for the company for each warranty purchased?

Solution 5

(a)

Mark scheme

- The 25th percentile of the standard normal distribution is $z = -0.6745$. The battery life span is normally distributed with mean 30 months and standard deviation 8 months, so the 25th percentile is $30 + 8(-0.6745) = 24.6$ months. Full credit requires: (1) indicating use of a normal distribution with mean 30 and sd 8, (2) a correct approach for finding the 25th percentile, (3) reporting the correct value 24.6 (or a value consistent with the setup).

Solution 5

(b)

Mark scheme

- $P(\text{life span} \leq 24 \text{ months}) = P\left(Z \leq \dfrac{24-30}{8}\right) = P(Z \leq -0.75) \approx 0.2266.$

  Full credit requires: (1) indicating the N(30,8) distribution, (2) specifying the correct event/boundary value/direction (life span $\leq$ 24 months), (3) reporting the correct probability ≈ 0.2266.

Solution 5

(c)

Mark scheme

- Expected gain per warranty $=(\$50)\times P(\text{life span}>24)+(-\$150)\times P(\text{life span}\le24) = (\$50)(0.7734)+(-\$150)(0.2266)\approx\$4.68$. Full credit requires: (1) an expected-value calculation using two probabilities that add to 1, correctly paired with the correct dollar outcomes ($+\$50$ for no replacement, $-\$150$ for replacement), (2) reporting the correct expected value (any value between \$4.00 and \$4.70 is accepted due to rounding, or any value between -\$150 and +\$50 consistent with the student's own work).

Question 3

2017 ▪ Q3

A grocery store purchases melons from two distributors, J and K. Distributor J provides melons from organic farms. The distribution of the diameters of the melons from Distributor J is approximately normal with mean 133 millimeters (mm) and standard deviation 5 mm.

(a) For a melon selected at random from Distributor J, what is the probability that the melon will have a diameter greater than 137 mm?

(b) Distributor K provides melons from nonorganic farms. The probability is 0.8413 that a melon selected at random from Distributor K will have a diameter greater than 137 mm. For all the melons at the grocery store, 70 percent of the melons are provided by Distributor J and 30 percent are provided by Distributor K.

For a melon selected at random from the grocery store, what is the probability that the melon will have a diameter greater than 137 mm?

Question 3

(c) Given that a melon selected at random from the grocery store has a diameter greater than 137 mm, what is the probability that the melon will be from Distributor J?

Solution 3

(a) Mark scheme

- Let X denote the diameter of a randomly selected Distributor J melon. X has an approximately normal distribution with mean 133 mm and standard deviation 5 mm, i.e. $X \sim N(133, 5)$. The z-score for a diameter of 137 mm is $z = (137 - 133)/5 = 4/5 = 0.8$. Therefore $P(X > 137) = P(Z > 0.8) = 1 - 0.7881 = 0.2119$.

Solution 3

- Full credit needs three components: (1) Normality and parameters – indicates use of a normal (or approximately normal) distribution and clearly identifies the correct mean (133) and standard deviation (5); (2) Boundary and direction – uses the correct boundary value ($x = 137$ or $z = 0.8$) and the correct direction (greater than / right tail); (3) Probability – reports the correct normal probability (0.2119) consistent with the setup in (1) and (2).

Solution 3

(b)

Mark scheme

- Define events: $J =$ melon is from Distributor J, $K =$ melon is from Distributor K, $G =$ melon diameter is greater than 137 mm. For a randomly selected melon from the grocery store, by the law of total probability: $P(G) = P(G \mid J)P(J) + P(G \mid K)P(K) = (0.2119)(0.7) + (0.8413)(0.3) = 0.1483 + 0.2524 = 0.4007$. (Equivalently, via a tree diagram: $P(G \text{ and } J) = 0.1483$ and $P(G \text{ and } K) = 0.2524$, which sum to 0.4007.) Full credit needs the probability computed correctly AND work shown using correct numerical values in a correct formula, tree diagram, or table.

Solution 3

(c) Mark scheme

- Using the events defined in part (b), the requested probability is $P(J \mid G) = \dfrac{P(J \text{ and } G)}{P(G)} = \dfrac{P(G \mid J)P(J)}{P(G)} = \dfrac{(0.2119)(0.7)}{0.4007} = \dfrac{0.1483}{0.4007} = 0.3701$. Full credit needs the probability computed correctly AND work shown that illustrates how it was found (the conditional-probability/Bayes formula with numerator and denominator identified, or the equivalent tree-diagram ratio). A correct answer that consistently follows from incorrect part (a)/(b) values (provided all probabilities stay between 0 and 1) still earns credit here.

Question 3

2014 ▪ Q3

Schools in a certain state receive funding based on the number of students who attend the school. To determine the number of students who attend a school, one school day is selected at random and the number of students in attendance that day is counted and used for funding purposes. The daily number of absences at High School A in the state is approximately normally distributed with mean of 120 students and standard deviation of 10.5 students.

(a) If more than 140 students are absent on the day the attendance count is taken for funding purposes, the school will lose some of its state funding in the subsequent year. Approximately what is the probability that High School A will lose some state funding?

Question 3

Each trial in the simulation consists of rolling three fair, six-sided dice, one die for each of the convention attendees. For each die, rolling a 1, 2, 3, or 4 represents selecting a man; rolling a 5 or 6 represents selecting a woman. After 1,000 trials, the number of times the dice indicate selecting 3 women is recorded.

Question 3

2014 ▪ Q3

Schools in a certain state receive funding based on the number of students who attend the school. To determine the number of students who attend a school, one school day is selected at random and the number of students in attendance that day is counted and used for funding purposes. The daily number of absences at High School A in the state is approximately normally distributed with mean of 120 students and standard deviation of 10.5 students.

(b) The principals' association in the state suggests that instead of choosing one day at random, the state should choose 3 days at random. With the suggested plan, High School A would lose some of its state funding in the subsequent year if the mean number of students absent for the 3 days is greater than 140. Would High School A be more likely, less likely, or equally likely to lose funding using the suggested plan compared to the plan described in part (a)? Justify your choice.

Question 3

2014 ▪ Q3

Schools in a certain state receive funding based on the number of students who attend the school. To determine the number of students who attend a school, one school day is selected at random and the number of students in attendance that day is counted and used for funding purposes. The daily number of absences at High School A in the state is approximately normally distributed with mean of 120 students and standard deviation of 10.5 students.

(c) A typical school week consists of the days Monday, Tuesday, Wednesday, Thursday, and Friday. The principal at High School A believes that the number of absences tends to be greater on Mondays and Fridays, and there is concern that the school will lose state funding if the attendance count occurs on a Monday or Friday. If one school day is chosen at random from each of 3 typical school weeks, what is the probability that none of the 3 days chosen is a Tuesday, Wednesday, or Thursday?

Solution 3

(a) Mark scheme

- (Q3 is graded holistically across parts (a)-(c) via a combination table into an overall 0-4 score; this leaf's marks approximate its share of the 4-point total.) The daily number of absences is approximately normal with mean 120 and standard deviation 10.5, so the z-score for 140 absences is $z = (140\text{-}120)/10.5 \approx 1.90$. The standard normal table/calculator gives $P(X \leq 140) = 0.9713$, so P(more than 140 absent, i.e. the school loses funding) $= 1 - 0.9713 = 0.0287$. Full credit needs three components: (1) indicates use of a normal distribution and clearly identifies the correct parameter values (showing correct components of a z-score calculation is sufficient), (2) uses a correct boundary value (140, 140.5, or 141 all acceptable), (3) reports the correct normal probability consistent with (1) and (2)

Solution 3

— OR reports a probability of 0.025 justified via the empirical rule with an acceptable boundary value. Partial credit for correctly providing only two of the three components, or for reporting an incorrect probability of 0.05 justified via the empirical rule with an acceptable boundary value. An inconsistency between calculation steps lowers the score for this part by one level.

Solution 3

(b) Mark scheme

- High School A would be LESS likely to lose state funding under the suggested 3-day plan. With a random sample of 3 days, the distribution of the sample mean number of absences has less variability than the distribution of a single day's absences, concentrating more narrowly around the mean of 120 and giving a smaller probability that it exceeds 140. Specifically, the standard deviation of the sample mean is $\sigma/\sqrt{n} = 10.5/\sqrt{3} \approx 6.062$, so the z-score for a sample mean of 140 is (140-120)/6.062 $\approx$ 3.30, giving a probability of 1 - 0.9995 = 0.0005 of losing funding under the suggested plan — less than the 0.0287 probability found in part (a). (An equivalent approach uses the total absences over 3 days: mean 3(120)=360, sd 3(6.062)$\approx$18.187, z=(420-360)/18.187$\approx$3.30, giving the same

Solution 3

probability; a response using this total-based approach is scored E if it gives the correct answer and references the distribution of the sample total with the correct mean and standard deviation.) Full credit needs the correct answer 'less likely' AND three components: (1) clearly references the distribution of the sample mean, (2) indicates the variability of that distribution is smaller (than for a single day), (3) indicates the distribution is centered at 120 — OR the correct answer 'less likely' AND both: correctly calculates the probability the sample mean exceeds 140 (arithmetic errors not penalized) AND correctly compares this probability to part (a)'s. Partial credit for 'less likely' with only two of the first three components, or for 'less likely' with a correctly-calculated exceedance probability but no explicit comparison to part (a). Incorrect if the response does not give the correct answer of 'less likely' with adequate support, including a correct answer given with no explanation or an incorrect explanation.

Solution 3

(c) Mark scheme

- For any one typical school week (Monday-Friday), P(the day selected is not Tuesday, Wednesday, or Thursday) $= 2/5 = 0.4$ (only Monday or Friday qualify). Because the days are selected independently across the three weeks, P(none of the 3 selected days is a Tuesday, Wednesday, or Thursday) $= (0.4)\hat{}3 = 0.064$. Full credit needs the response to correctly calculate the probability AND show sufficient supporting work. Partial credit for reporting the correct probability (0.064) with no work shown or insufficient work shown, or for using the multiplication rule involving the three weekly selections but doing so incorrectly and/or with an incorrect single-week probability of not selecting a Tuesday, Wednesday, or Thursday.