Past-paper walk-through

Parameters for a Binomial Distribution ▪ 4.11

eXercise

Questions in this set

- 2026 Q3
- 2025 Q3
- 2022 Q3
- 2021 Q3
- 2019 Q3
- 2018 Q3
- 2016 Q4

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 3

2025 ▪ Q3

Ms. Fey is a manager at a restaurant. To improve the dining experience for her customers, she uses a digital music service to create a playlist of songs that will be played in the restaurant. The playlist contains 1,000 songs and consists of four different types of music in the following quantities: 200 country songs, 400 pop songs, 100 rock songs, and 300 jazz songs. The digital music service will select songs at random from the playlist to be played in the restaurant. Any song can be replayed at any time.

(a)(i) Suppose one song is selected at random to be played. What is the probability that the song is a rock song? Show your work.

(a)(ii) Suppose two songs are selected at random to be played. What is the probability that both songs are rock songs? Show your work.

Question 3

2025 ▪ Q3

Ms. Fey is a manager at a restaurant. To improve the dining experience for her customers, she uses a digital music service to create a playlist of songs that will be played in the restaurant. The playlist contains 1,000 songs and consists of four different types of music in the following quantities: 200 country songs, 400 pop songs, 100 rock songs, and 300 jazz songs. The digital music service will select songs at random from the playlist to be played in the restaurant. Any song can be replayed at any time.

(b)(i) In every one-hour period, 20 songs will be played at random and any song can be replayed at any time. Ms. Fey is interested in how many rock songs will be played in a typical one-hour period.

Define the random variable of interest to Ms. Fey, and state how the random variable is distributed.

Question 3

(b)(ii) What is the expected value for the random variable in part B (i)? Show your work.

Question 3

2025 ▪ Q3

Ms. Fey is a manager at a restaurant. To improve the dining experience for her customers, she uses a digital music service to create a playlist of songs that will be played in the restaurant. The playlist contains 1,000 songs and consists of four different types of music in the following quantities: 200 country songs, 400 pop songs, 100 rock songs, and 300 jazz songs. The digital music service will select songs at random from the playlist to be played in the restaurant. Any song can be replayed at any time.

(c)(i) Recall that in every one-hour period, 20 songs will be played at random and any song can be replayed at any time.

Determine the probability that 4 or more rock songs in a particular one-hour period will be played. Show your work.

Question 3

(c)(ii) Suppose 4 rock songs are played during a particular one-hour period. Does this provide strong evidence that the song selection process was not truly random? Justify your answer without performing an inference procedure.

Solution 3

(a)(i)

Mark scheme

- P(Rock Song) $= 100/1{,}000 = 0.10$. Full credit needs both: (1) the response calculates the correct probability (0.10), (2) the response shows supporting work for it (the 100/1000 ratio).

Solution 3

(a)(ii) Mark scheme

- P(Both Rock Songs) $= (0.10)(0.10) = 0.01$, using independence of the two song selections (any song can be replayed / selections are treated as independent). This leaf's credit is bundled with 3(a)(i) in the official scoring (Part A 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, independence-consistent answer of 0.01 with supporting work is still expected and reinforces credit for part (a). Note: a response that instead samples without replacement, e.g. $(100/1000)(99/999) \approx 0.0099$, is also acceptable if it shows valid supporting work, though it does not reflect the 'each song is replayed randomly' independence framing given in the stem.

Solution 3

(b)(i) Mark scheme

- Let the random variable X represent the number of rock songs (out of the 20 songs played) in one hour. Since any song can be replayed at any time, each of the 20 songs played is independently a rock song with probability $100/1{,}000 = 0.10$, so X has a binomial distribution with $n = 20$ independent trials and probability of success $p = 0.10$ for each trial, i.e. $X \sim B(20, 0.10)$. Full credit needs at least 3 of 4 components: (1) defines the random variable as the number of rock songs played in one hour (must include both 'number of rock songs' AND 'in one hour'/'out of 20 songs'), (2) describes the distribution as binomial, (3) states $n = 20$ and $p = 0.10$ (may appear in part (b)(i) or (b)(ii)), (4) [see 3(b)(ii)]. Stating songs are 'distributed randomly' is not itself a distribution

Solution 3

statement and should be treated as extraneous; if the response names a non-binomial distribution, this part cannot score essentially correct.

Solution 3

(b)(ii)

Mark scheme

- $E(X) = np = 20(0.10) = 2$ rock songs. This leaf's credit is bundled with 3(b)(i) in the official scoring (Part B of the SG scores both sub-parts together), so it carries 0 of this question's marks on its own — but the correct expected value (2) with the np calculation shown (e.g. $20(0.10) = 2$) is still expected to complete credit for part (b).

Solution 3

(c)(i)

Mark scheme

- $\mathsf{P}(\mathsf{X} \geq 4) = 1 - \mathsf{P}(\mathsf{X} \leq 3) = 1 - [\binom{20}{0}(0.10)^0(0.90)^{20} + \binom{20}{1}(0.10)^1(0.90)^{19} + \binom{20}{2}(0.10)^2(0.90)^{18} + \binom{20}{3}(0.10)^3(0.90)^{17}] = 1{-}0.867 = 0.133. Fullcreditneeds : (1) thecorrectprobability 0.133, (2) supportingwork—anexplicitbinomial - probabilityformula, alabeledbargraphofbinomialprobabilities, orclearly - labeledcalculatornotationsuchas'1{-}binomcdf(n = 20, p = 0.10, upperbound = 3)'(anunlabeled' binomcdf(20, 0.10, 3)' isnotsufficientsupport). Anarithmetic/transcriptionerro$

Solution 3

(c)(ii) Mark scheme

- No, this does not provide strong evidence that the song-selection process was not truly random. The probability that 4 or more rock songs would be played in a particular one-hour period is 0.133 (from part (c)(i)), which is high enough to be reasonably attributed to random chance alone — it is not small enough to constitute strong evidence against the process being truly random. Full credit needs: (1) indicates there is not a strong reason to believe the selection process was not truly random (i.e., answers 'No'), (2) an explanation that correctly links the (c)(i) probability to that decision (0.133 is not a small/surprising probability). Note: a response that instead argues 'Yes' (there IS strong evidence) can still earn full credit here if it is adequately supported by a correctly-computed alternative

Solution 3

probability or by the standard deviation of the binomial distribution ($\approx$1.342) combined with an explanation based on 4 being within two standard deviations of the mean of 2.

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 $P(A < 0.5) = P\left(Z < \dfrac{0.5 - 0.6}{0.04} = -2.5\right) \approx 0.0062$. 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 3

2021 ▪ Q3

To increase morale among employees, a company began a program in which one employee is randomly selected each week to receive a gift card. Each of the company's 200 employees is equally likely to be selected each week, and the same employee could be selected more than once. Each week's selection is independent from every other week.

(a)(i) Consider the probability that a particular employee receives at least one gift card in a 52-week year.

Define the random variable of interest and state how the random variable is distributed.

(a)(ii) Determine the probability that a particular employee receives at least one gift card in a 52-week year. Show your work.

(b) Calculate and interpret the expected value for the number of gift cards a particular employee will receive in a 52-week year. Show your work.

Question 3

(c) Suppose that Agatha, an employee at the company, never receives a gift card for an entire 52-week year. Based on her experience, does Agatha have a strong argument that the selection process was not truly random? Explain your answer.

Solution 3

(a)(i)

Mark scheme

- Let $X =$ the number of gift cards a particular employee receives in a 52-week year. Each week the employee has probability $1/200 = 0.005$ of being selected, independently from week to week, so X follows a binomial distribution with $n = 52$ trials and probability of success $p = 0.005$: $X \sim B(52, 0.005)$. Credit requires defining the random variable in context AND correctly naming/stating the binomial distribution with $n = 52, p = 0.005$.

Solution 3

(a)(ii)

Mark scheme

- $P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{52}{0}(0.005)^0(0.995)^{52} = 1 - 0.7705 = 0.2295$ $(\approx 0.230). Credit requires identifying the event as X \geq 1$ (or its complement $X = 0$) with correct boundary/direction AND supporting work leading to the correct probability 0.2295 (or a value consistent with a correctly identified binomial setup).

Solution 3

(b)

Mark scheme

- Expected value $E(X) = np = 52(0.005) = 0.26$ gift cards. Interpretation: if the weekly random selection process were repeated for a very large number of 52-week years, each employee would receive about 0.26 gift cards per year on average (about one gift card every four years). Credit requires (1) the correct expected-value calculation with supporting work, AND (2) a reasonable interpretation mentioning at least two of: repeating the process long-run, the concept of an average/mean, and the context of receiving gift cards.

Solution 3

(c) Mark scheme

- No, Agatha's experience is not strong evidence the selection process wasn't truly random. It is quite likely for a particular employee to receive zero gift cards in an entire 52-week year: $P(X = 0) = (0.995)^{52} \approx 0.7705$ (about 77

Question 3

2019 ▪ Q3

A medical researcher surveyed a large group of men and women about whether they take medicine as prescribed. The responses were categorized as never, sometimes, or always. The relative frequency of each category is shown in the table.

	Never	**Sometimes**	**Always**	**Total**
Men	0.0564	0.2016	0.2120	0.4700
Women	0.0636	0.1384	0.3280	0.5300
Total	0.1200	0.3400	0.5400	1.0000

Question 3

(a)(i) One person from those surveyed will be selected at random.
What is the probability that the person selected will be someone whose response is never and who is a woman?
(a)(ii) What is the probability that the person selected will be someone whose response is never or who is a woman?
(a)(iii) What is the probability that the person selected will be someone whose response is never given that the person is a woman?
(b) For the people surveyed, are the events of being a person whose response is never and being a woman independent? Justify your answer.

Question 3

(c) Assume that, in a large population, the probability that a person will always take medicine as prescribed is 0.54. If 5 people are selected at random from the population, what is the probability that at least 4 of the people selected will always take medicine as prescribed? Support your answer.

Question 3

The responses were categorized as never, sometimes, or always. The relative frequency of each category shown in the table.

	Never	Sometimes	Always	Total
Men	0.0564	0.2016	0.2120	0.4700
Women	0.0636	0.1384	0.3280	0.5300
Total	0.1200	0.3400	0.5400	1.0000

Solution 3

(a)(i)

Mark scheme

- $P(\text{never and woman}) = 0.0636$, read directly as the joint relative frequency from the two-way table.

(a)(ii)

Mark scheme

- $P(\text{never or woman}) = P(\text{never}) + P(\text{woman}) - P(\text{never and woman}) = 0.12 + 0.53 - 0.0636 = 0.5864$. (Equivalently, sum the relative frequencies for all cells that are 'never' or 'woman': $0.0564 + 0.0636 + 0.1384 + 0.3280 = 0.5864$, or $0.0564 + 0.53 = 0.5864$.)

Solution 3

(a)(iii)

Mark scheme

- $P(\text{never} \mid \text{woman}) = \dfrac{P(\text{never and woman})}{P(\text{woman})} = \dfrac{0.0636}{0.53} = 0.12.$

Solution 3

(b)

Mark scheme

- Yes, the event of being a person who responds 'never' is independent of the event of being a woman, because $P(\text{never} \mid \text{woman}) = 0.12 = P(\text{never})$ – the conditional probability equals the unconditional (marginal) probability, using numbers from the table. (An equally valid justification: $P(\text{woman and never}) = 0.0636$ is the same as $P(\text{woman}) \times P(\text{never}) = (0.53)(0.12) = 0.0636$.) Must state the events are independent AND justify with actual numbers from the table.

Solution 3

(c) Mark scheme

- Let $X =$ the number of people (out of 5 randomly selected) who always take their medicine as prescribed. X has a binomial distribution with $n = 5$ and $p = 0.54$. $P(X \geq 4) = \binom{5}{4}(0.54)^4(0.46)^1 + \binom{5}{5}(0.54)^5(0.46)^0 \approx 0.19557 + 0.04592 \approx 0.24149$. Full credit requires: (1) clearly identifying the binomial distribution with $n = 5, p = 0.54$, (2) correctly identifying the boundary value and direction (at least 4, i.e. $X \geq 4$), and (3) reporting the correct probability ≈ 0.241 (accept values consistent with a correct binomial setup, e.g. 0.24-0.2415). Note: a normal approximation is NOT appropriate here since $np = 5 \times 0.54 = 2.7 < 5$.

Question 3

2018 ▪ Q3

Approximately 3.5 percent of all children born in a certain region are from multiple births (that is, twins, triplets, etc.). Of the children born in the region who are from multiple births, 22 percent are left-handed. Of the children born in the region who are from single births, 11 percent are left-handed.

(a) What is the probability that a randomly selected child born in the region is left-handed?

(b) What is the probability that a randomly selected child born in the region is a child from a multiple birth, given that the child selected is left-handed?

(c) A random sample of 20 children born in the region will be selected. What is the probability that the sample will have at least 3 children who are left-handed?

Solution 3

(a)

Mark scheme

- Let L = left-handed, M = multiple birth, S = single birth. By the law of total probability: $P(L) = P(M)P(L \mid M) + P(S)P(L \mid S) = (0.035)(0.22) + (0.965)(0.11) = 0.0077 + 0.10615 = 0.11385$. Full credit requires the probability computed correctly AND work shown using correct numerical values via a formula, tree diagram, or other appropriate strategy — in particular, summing the results of two multiplications.

Solution 3

(b) Mark scheme

- Using the definition of conditional probability together with the part-(a) result $P(L) = 0.11385$: $P(M \mid L) = \dfrac{P(L \text{ and } M)}{P(L)} = \dfrac{(0.035)(0.22)}{0.11385} = \dfrac{0.0077}{0.11385} \approx 0.0676.$
  Full credit requires the probability computed correctly with work shown that includes appropriate numerical values for both the numerator and the denominator (values carried over from an incorrect part (a) are still acceptable as 'appropriate').

Solution 3

(c) Mark scheme

- Let $X =$ number of left-handed children in a random sample of 20 children from the region. From part (a),
  X is Binomial with $n = 20$ and $p = 0.11385$. Then $P(X \geq 3) = 1 - P(X \leq 2) = 1 - \left[\binom{20}{0}(0.11385)^0(0.88615)^{20} + \binom{20}{1}(0.11385)^1(0.88615)^{19} + \binom{20}{2}(0.11385)^2(0.88615)^{18}\right] \approx$ $1 - 0.598 \approx 0.402$. Full credit (5 components) requires: using a calculation based on the binomial distribution; specifying appropriate values for n and p; using the correct endpoint (X˜22642, i.e. the complement of X˜22653, not X˜22643); using the correct direction (at-least-3 via the complement); and correctly calculating a binomial probability consistent with earlier work. Note: a normal approximation is NOT appropriate here because $np = 20 \times 0.11385 = 2.277 < 5$; a response using

Solution 3

the normal approximation can score at most partial credit, and only then if it shows a correct mean/SD from the binomial parameters, a clear boundary/direction (z-score or diagram), and a correctly computed probability.

Question 4

2016 ▪ Q4

A company manufactures model rockets that require igniters to launch. Once an igniter is used to launch a rocket, the igniter cannot be reused. Sometimes an igniter fails to operate correctly, and the rocket does not launch. The company estimates that the overall failure rate, defined as the percent of all igniters that fail to operate correctly, is 15 percent.

A company engineer develops a new igniter, called the super igniter, with the intent of lowering the failure rate. To test the performance of the super igniters, the engineer uses the following process.

Step 1: One super igniter is selected at random and used in a rocket.

Step 2: If the rocket launches, another super igniter is selected at random and used in a rocket.

Question 4

Step 2 is repeated until the process stops. The process stops when a super igniter fails to operate correctly or 32 super igniters have successfully launched rockets, whichever comes first. Assume that super igniter failures are independent.

(a) If the failure rate of the super igniters is 15 percent, what is the probability that the first 30 super igniters selected using the testing process successfully launch rockets?

(b) Given that the first 30 super igniters successfully launch rockets, what is the probability that the first failure occurs on the thirty-first or the thirty-second super igniter tested if the failure rate of the super igniters is 15 percent?

(c) Given that the first 30 super igniters successfully launch rockets, is it reasonable to believe that the failure rate of the super igniters is less than 15 percent? Explain.

Solution 4

(a)

Mark scheme

- If the failure rate is 15

Solution 4

(b)

Mark scheme

- Given the first 30 succeed, the probability the first failure occurs on the 31st trial is 0.15, and the probability it occurs on the 32nd trial (31st succeeds, 32nd fails) is $(0.85)(0.15) = 0.1275$. So $P(\text{first failure on the } 31^{st} \text{ or } 32^{nd}) = 0.15 + 0.1275 = 0.2775$. (This is equivalent to asking for the probability the first failure occurs on the first or second of two new trials, since the process is memoryless.) Full credit requires both the correct probability AND correct justification.

Solution 4

(c) Mark scheme

- Yes, it is reasonable to believe the failure rate of the super igniters is less than 15