Mathematical Methods Units 3 & 4

Chapter 12 — The binomial distribution

Aboriginal and Torres Strait Islander content has been excluded as accuracy cannot be guaranteed, and it is better to be respectful than cover the entire curriculum inaccurately.

Contents

Chapter 12 The binomial distribution — 12A Bernoulli trials and the binomial distribution

Theory

Bernoulli trials. A **Bernoulli trial** is a random experiment whose only two possible outcomes are labelled *success* and *failure*. A run of such trials forms a **Bernoulli sequence** when

- each trial has the same probability of success p (so failure has probability $1 - p$), and
- the trials are **independent**.

The binomial distribution. Let X be the number of successes in n independent Bernoulli trials, each with success probability p. Then X has a **binomial distribution**, written

$$X \sim Bi(n, p),$$

and its probability of exactly k successes is

$$Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, k = 0, 1, 2, \ldots, n,$$

where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ counts the number of ways of choosing which k of the n trials are the successes among the n trials.

When is a model binomial? Check all four conditions: a **fixed** number of trials n; **two** outcomes per trial; a **constant** success probability p; and **independent** trials. Sampling *without* replacement (so p changes) is **not** binomial.

Useful probabilities.

- "At least one": $Pr(X \geq 1) = 1 - Pr(X = 0)$.
- "At most / at least k": add the required terms, e.g. $Pr(X \leq 2) = Pr(0) + Pr(1) + Pr(2)$.
- Conditional: $Pr(X = a / X \geq b) = \frac{Pr(X = a)}{Pr(X \geq b)}$ (when $a \geq b$).

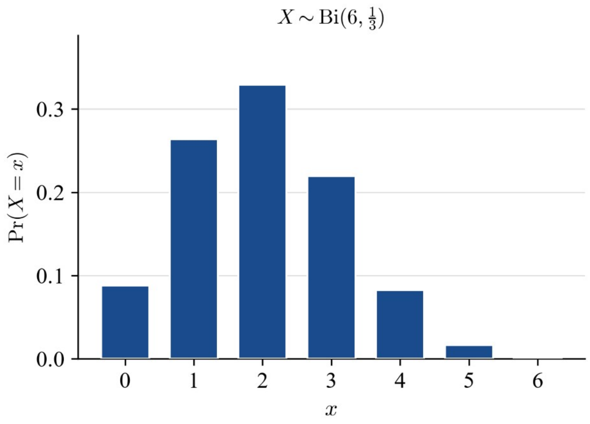
Give probabilities as **exact fractions** unless a decimal is asked for.

Worked examples

Example 1 — scaffolded

The probability that a biased spinner lands on red is $1/3$. It is spun 6 times. Let X be the number of reds. Find $Pr(X = 2)$.

Working	Comment
Each spin is red (success, $p = 1/3$) or not; 6 independent spins, so $X \sim Bi(6, 1/3)$.	Fixed $n = 6$, two outcomes, constant p, independent — binomial.
$Pr(X = 2) = \binom{6}{2} (1/3)^2 (2/3)^4$	Apply the formula with $k = 2$.
$= 15 \cdot \frac{1}{9} \cdot \frac{16}{81} = \frac{240}{729}$	$\binom{6}{2} = 15$; multiply.
$\therefore Pr(X = 2) = \frac{80}{243}$	Simplify.



Example 2 — scaffolded

A archer hits the bullseye with probability $2/5$ on each shot. In 5 independent shots, let X be the number of bullseyes. Find $Pr(X \geq 4)$.

Working	Comment
$X \sim Bi(5, 2/5)$; $Pr(X \geq 4) = Pr(X = 4) + Pr(X = 5)$.	"At least 4" means 4 or 5.
$Pr(X = 4) = \binom{5}{4}(2/5)^4(3/5) = 5 \cdot \frac{16}{625} \cdot \frac{3}{5} = \frac{240}{3125}$	$k = 4$.
$Pr(X = 5) = (2/5)^5 = \frac{32}{3125}$	$k = 5$: all successes.
$\therefore Pr(X \geq 4) = \frac{240 + 32}{3125} = \frac{272}{3125}$	Add.

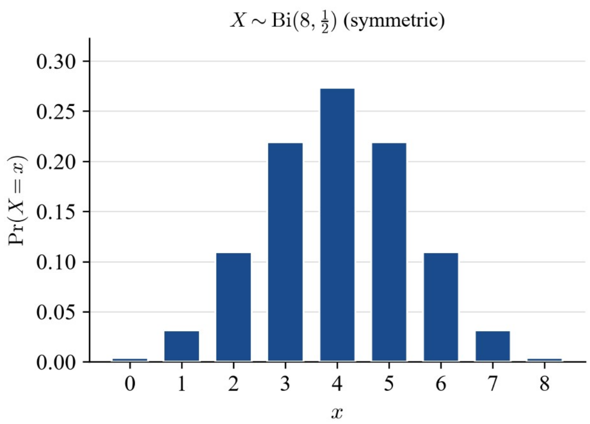
Example 3 — unscaffolded

A fair coin is tossed 8 times. Find the probability of exactly 3 heads.

Let X be the number of heads, $X \sim Bi(8, 1/2)$.

$$Pr(X = 3) = \binom{8}{3}(1/2)^3(1/2)^5 = \binom{8}{3}(1/2)^8 = \frac{56}{256}.$$

$$\therefore Pr(X = 3) = \frac{7}{32}.$$



Example 4 — unscaffolded

A fair coin is tossed 4 times. Find the probability of getting **at least one** head.

Let X be the number of heads, $X \sim Bi(4, 1/2)$. Using the complement,

$Pr(X \geq 1) = 1 - Pr(X = 0) = 1 - (1/2)^4 = 1 - \frac{1}{16}.$

$\therefore Pr(X \geq 1) = \frac{15}{16}.$

Practice questions

Scaffolded

Q1. $X \sim Bi(7, 1/2)$. (a) Write the formula for $Pr(X = 4)$; (b) evaluate $\binom{7}{4}$; (c) hence find $Pr(X = 4)$.

Q2. A machine produces components, each defective with probability $1/5$, independently. A sample of 10 is taken. Let X be the number defective. (a) State the distribution of X; (b) find $Pr(X = 2)$.

Q3. $X \sim Bi(5, 1/3)$. (a) Find $Pr(X = 0)$; (b) find $Pr(X = 1)$; (c) hence find $Pr(X \leq 1)$.

Q4. A fair die is rolled 6 times; a "success" is rolling a 1 or a 2. Let X be the number of successes. (a) State p and the distribution of X; (b) find $Pr(X \geq 5)$.

Q5. $X \sim Bi(4, 3/4)$. (a) Find $Pr(X = 4)$; (b) find $Pr(X = 0)$; (c) which of $Pr(X = 4)$ and $Pr(X = 0)$ is larger, and why?

Q6. $X \sim Bi(5, 2/5)$. Find $Pr(X = 2)$.

Unscaffolded

Give each probability as an exact fraction.

Q7. A fair coin is tossed 10 times. Find the probability of exactly 5 heads.

Q8. For $X \sim Bi(6, 1/3)$, find $Pr(X \geq 1)$.

Q9. $X \sim Bi(8, 3/4)$. Find $Pr(X = 6)$.

Q10. $X \sim Bi(5, 1/2)$. Find $Pr(X \leq 2)$.

Q11. $X \sim Bi(7, 2/5)$. Find $Pr(X = 3)$.

Q12. $X \sim Bi(4, 1/3)$. Find $Pr(X \geq 3)$.

Q13. $X \sim Bi(6, 1/2)$. Find $Pr(2 \leq X \leq 4)$.

Q14. $X \sim Bi(10, 1/10)$. Find $Pr(X = 1)$.

Q15. $X \sim Bi(5, 4/5)$. Find $Pr(X = 5)$.

Q16. $X \sim Bi(6, 1/4)$. Show that $Pr(X \geq 5)$ can be written in the form $\frac{a}{4^6}$ and state a.

Q17. $X \sim Bi(9, 1/3)$. Find $Pr(X = 3)$.

Q18. A fair coin is tossed 4 times. Given that at least one head occurs, find the probability of exactly two heads.

Q19. $X \sim Bi(5, 1/2)$. Find $Pr(X \geq 3)$.

Q20. For each situation, state with a reason whether X has a binomial distribution. (a) X = the number of sixes in 10 rolls of a fair die. (b) X = the number of aces when 5 cards are dealt from a standard deck (no replacement). (c) X = the number of heads in tosses of a coin until the first tail.

Q21. $X \sim Bi(6, 1/3)$. Given that $X \leq 3$, find the probability that $X = 2$.

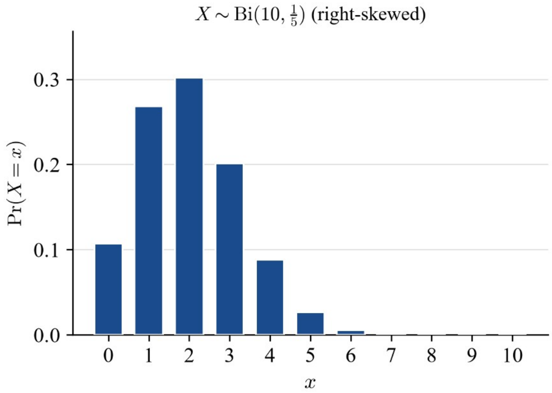
Answers

Q1. (c) $\frac{35}{128}$. **Q2.** (a) $Bi(10,1/5)$; (b) $\frac{589824}{1953125}$. **Q3.** $\frac{112}{243}$. **Q4.** (a) $p=1/3$, $Bi(6,1/3)$; (b) $\frac{13}{729}$. **Q5.** (a) $\frac{81}{256}$; (b) $\frac{1}{256}$. **Q6.** $\frac{216}{625}$. **Q7.** $\frac{63}{256}$. **Q8.** $\frac{665}{729}$. **Q9.** $\frac{5103}{16384}$. **Q10.** $\frac{1}{2}$. **Q11.** $\frac{4536}{15625}$. **Q12.** $\frac{1}{9}$. **Q13.** $\frac{25}{32}$. **Q14.** $\frac{387420489}{1000000000}$. **Q15.** $\frac{1024}{3125}$. **Q16.** $a=19$. **Q17.** $\frac{1792}{6561}$. **Q18.** $\frac{2}{5}$. **Q19.** $\frac{1}{2}$. **Q20.** (a) yes; (b) no; (c) no. **Q21.** $\frac{15}{41}$.

Fully-worked solutions

Q1. $X\sim Bi(7,1/2)$. (a) $Pr(X=4)=\binom{7}{4}(1/2)^4(1/2)^3=\binom{7}{4}(1/2)^7$. (b) $\binom{7}{4}=35$. (c) $Pr(X=4)=\frac{35}{128}$.

Q2. $X\sim Bi(10,1/5)$. $Pr(X=2)=\binom{10}{2}(1/5)^2(4/5)^8=45\cdot\frac{1}{25}\cdot\frac{65536}{390625}=\frac{589824}{1953125}$.



Q3. $X\sim Bi(5,1/3)$. $Pr(X=0)=(2/3)^5=\frac{32}{243}$; $Pr(X=1)=\binom{5}{1}(1/3)(2/3)^4=5\cdot 1/3\cdot\frac{16}{81}=\frac{80}{243}$;
$Pr(X\le 1)=\frac{32+80}{243}=\frac{112}{243}$.

Q4. A 1 or 2 has probability $p=2/6=1/3$, so $X\sim Bi(6,1/3)$.
$Pr(X\ge 5)=Pr(X=5)+Pr(X=6)=\binom{6}{5}(1/3)^5(2/3)+(1/3)^6=\frac{12}{729}+\frac{1}{729}=\frac{13}{729}$.

Q5. $X\sim Bi(4,3/4)$. (a) $Pr(X=4)=(3/4)^4=\frac{81}{256}$. (b) $Pr(X=0)=(1/4)^4=\frac{1}{256}$. (c) $Pr(X=4)$ is much larger because success is likely ($p=3/4$), so all four succeeding is far more probable than none.

Q6. $Pr(X=2)=\binom{5}{2}(2/5)^2(3/5)^3=10\cdot\frac{4}{25}\cdot\frac{27}{125}=\frac{216}{625}$.

Q7. $X\sim Bi(10,1/2)$. $Pr(X=5)=\binom{10}{5}(1/2)^{10}=\frac{252}{1024}=\frac{63}{256}$.

Q8. $Pr(X\ge 1)=1-Pr(X=0)=1-(2/3)^6=1-\frac{64}{729}=\frac{665}{729}$.

Q9. $Pr(X=6)=\binom{8}{6}(3/4)^6(1/4)^2=28\cdot\frac{729}{4096}\cdot\frac{1}{16}=\frac{5103}{16384}$.

Q10. $Pr(X \le 2) = (1/2)^5\left[\binom{5}{0}+\binom{5}{1}+\binom{5}{2}\right] = \frac{1+5+10}{32} = \frac{16}{32} = \frac{1}{2}.$

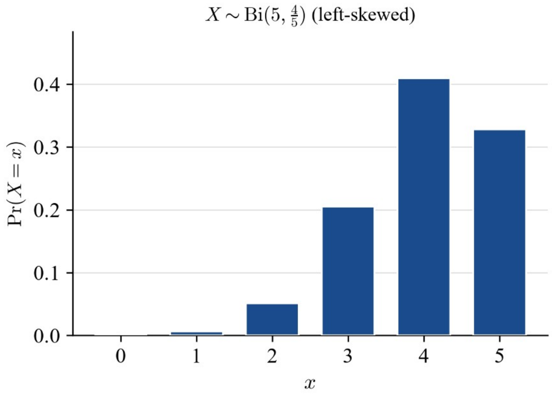
Q11. $Pr(X=3) = \binom{7}{3}(2/5)^3(3/5)^4 = 35 \cdot \frac{8}{125} \cdot \frac{81}{625} = \frac{4536}{15625}.$

Q12. $Pr(X \ge 3) = \binom{4}{3}(1/3)^3(2/3) + (1/3)^4 = \frac{8}{81} + \frac{1}{81} = \frac{9}{81} = \frac{1}{9}.$

Q13. $Pr(2 \le X \le 4) = (1/2)^6\left[\binom{6}{2}+\binom{6}{3}+\binom{6}{4}\right] = \frac{15+20+15}{64} = \frac{50}{64} = \frac{25}{32}.$

Q14. $Pr(X=1) = \binom{10}{1}(1/10)(9/10)^9 = \frac{9^9}{10^9} = \frac{387420489}{1000000000}.$

Q15. $Pr(X=5) = (4/5)^5 = \frac{1024}{3125}.$



Q16. $Pr(X \ge 5) = \binom{6}{5}(1/4)^5(3/4) + (1/4)^6 = \frac{6 \cdot 3}{4^6} + \frac{1}{4^6} = \frac{18+1}{4^6} = \frac{19}{4^6}$. $\therefore a = 19$ (and $\frac{19}{4096}$).

Q17. $Pr(X=3) = \binom{9}{3}(1/3)^3(2/3)^6 = 84 \cdot \frac{1}{27} \cdot \frac{64}{729} = \frac{5376}{19683} = \frac{1792}{6561}.$

Q18. $X \sim Bi(4, 1/2)$. $Pr(X=2 / X \ge 1) = \frac{Pr(X=2)}{Pr(X \ge 1)}$. $Pr(X=2) = \binom{4}{2}(1/2)^4 = \frac{6}{16} = \frac{3}{8}$;
$Pr(X \ge 1) = 1-(1/2)^4 = \frac{15}{16}$. So the probability $= \frac{3/8}{15/16} = \frac{3}{8} \cdot \frac{16}{15} = \frac{2}{5}.$

Q19. $Pr(X \ge 3) = (1/2)^5\left[\binom{5}{3}+\binom{5}{4}+\binom{5}{5}\right] = \frac{10+5+1}{32} = \frac{16}{32} = \frac{1}{2}.$

Q20. (a) **Yes** — $n=10$ fixed, two outcomes (six / not six), constant $p=1/6$, independent rolls. (b) **No** — dealing without replacement changes p from card to card, so the trials are not independent and p is not constant. (c) **No** — the number of trials is not fixed (it depends on when the first tail appears).

Q21. $Pr(X=2 / X \le 3) = \frac{Pr(X=2)}{Pr(X \le 3)}$. $Pr(X=2) = \binom{6}{2}(1/3)^2(2/3)^4 = \frac{80}{243}.$
$Pr(X \le 3) = (1/3)^6\left[\binom{6}{0}2^6+\binom{6}{1}2^5+\binom{6}{2}2^4+\binom{6}{3}2^3\right] = \frac{64+192+240+160}{729} = \frac{656}{729}$. So the probability
$= \frac{80/243}{656/729} = \frac{80}{243} \cdot \frac{729}{656} = \frac{240}{656} = \frac{15}{41}.$

Theory

If X counts the number of successes in a **Bernoulli sequence** of n independent trials, each with probability of success p, then X has a **binomial distribution**, written $X \sim \mathrm{Bi}(n, p)$, with

$$Pr(X=k)=\binom{n}{k}p^k(1-p)^{n-k}, k=0,1,2,\ldots,n.$$

The graph. The probability distribution is drawn as a bar graph (the height of the bar at $x=k$ is $Pr(X=k)$). Its **shape** depends on p:

- $p=\frac{1}{2}$ — the graph is **symmetric** about the mean;
- $p<\frac{1}{2}$ — the graph is **positively skewed** (longer tail to the right);
- $p>\frac{1}{2}$ — the graph is **negatively skewed** (longer tail to the left).

Increasing n makes the graph more bell-shaped. The **mode** (most likely value) is the value of k with the tallest bar, and it lies close to the mean.

Expectation, variance and standard deviation. For $X \sim \mathrm{Bi}(n, p)$ (writing $q=1-p$):

$$E(X)=np, \mathrm{Var}(X)=np(1-p)=npq, \mathrm{sd}(X)=\sqrt{np(1-p)}.$$

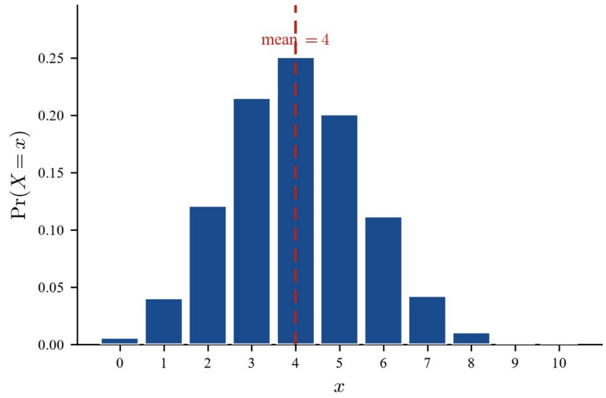
These follow from the general definitions $E(X)=\sum x\,Pr(X=x)$ and $\mathrm{Var}(X)=E(X^2)-[E(X)]^2$ (the proofs are in 12D). Note the variance is **largest when** $p=\frac{1}{2}$ and shrinks as p approaches 0 or 1. Take care to distinguish the **variance** $np(1-p)$ from the **standard deviation** $\sqrt{np(1-p)}$ — the standard deviation is what has the same units as X.

Worked examples

Example 1 — scaffolded

For $X \sim \mathrm{Bi}(10, 2/5)$, find the mean, variance and standard deviation, and describe the shape of the graph.

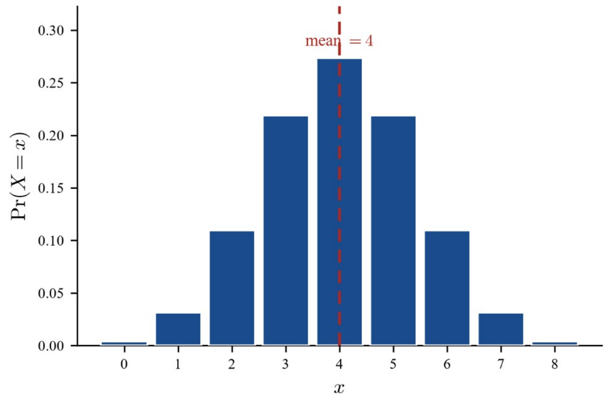
Working	Comment
$n=10$, $p=\frac{2}{5}$, so $q=1-\frac{2}{5}=\frac{3}{5}$.	Identify the parameters.
$E(X)=np=10\times\frac{2}{5}=4$.	Mean $=np$.
$\mathrm{Var}(X)=np(1-p)=10\times\frac{2}{5}\times\frac{3}{5}=\frac{12}{5}$.	Variance $=np(1-p)$.
$\mathrm{sd}(X)=\sqrt{\frac{12}{5}}=\frac{2\sqrt{15}}{5}$.	Standard deviation $=\sqrt{\mathrm{Var}(X)}$; keep it exact.
Since $p=\frac{2}{5}<\frac{1}{2}$, the graph is **positively skewed**.	The sign of $p-1/2$ fixes the skew.



Example 2 — scaffolded

For $X \sim \mathrm{Bi}(8, 1/2)$, find the mean, variance and standard deviation, state the mode, and describe the graph.

Working	Comment
$n=8,\ p=\frac{1}{2}$.	Parameters.
$E(X)=8\times\frac{1}{2}=4$.	Mean $= n\,p$.
$\mathrm{Var}(X)=8\times\frac{1}{2}\times\frac{1}{2}=2$; $\mathrm{sd}(X)=\sqrt{2}$.	Variance $= n\,p(1-p)$.
The graph is **symmetric** (since $p=\frac{1}{2}$), so the tallest bar is at the mean: the **mode is** 4.	For $p=1/2$ the mode equals the mean.



Example 3 — unscaffolded

For $X \sim \mathrm{Bi}(6, 1/3)$, find the mean, variance and standard deviation.

$$E(X)=n\,p=6\times\frac{1}{3}=2.$$

$$\mathrm{Var}(X)=n\,p(1-p)=6\times\frac{1}{3}\times\frac{2}{3}=\frac{4}{3}.$$

$$\mathrm{sd}(X)=\sqrt{\frac{4}{3}}=\frac{2}{\sqrt{3}}=\frac{2\sqrt{3}}{3}.$$

$\therefore$ mean 2, variance $\frac{4}{3}$, standard deviation $\frac{2\sqrt{3}}{3}$.

Example 4 — unscaffolded

A binomial random variable $X \sim \mathrm{Bi}(n, p)$ has mean 15 and variance $\frac{45}{4}$. Find n and p, and hence the standard deviation.

$np = 15$ and $np(1-p) = \frac{45}{4}$.

Dividing: $\frac{np(1-p)}{np} = 1 - p = \frac{45/4}{15} = \frac{3}{4}$, so $p = \frac{1}{4}$.

Then $n = \frac{15}{p} = \frac{15}{1/4} = 60$.

$\mathrm{sd}(X) = \sqrt{\frac{45}{4}} = \frac{3\sqrt{5}}{2}$.

$\therefore n = 60$, $p = \frac{1}{4}$, standard deviation $\frac{3\sqrt{5}}{2}$.

Practice questions

Scaffolded

Q1. For $X \sim \mathrm{Bi}(12, 1/4)$: (a) state n, p and q; (b) find the mean; (c) find the variance; (d) find the standard deviation.

Q2. For $X \sim \mathrm{Bi}(20, 1/2)$: (a) find the mean and variance; (b) find the standard deviation; (c) is the graph symmetric, positively skewed or negatively skewed?

Q3. For $X \sim \mathrm{Bi}(18, 1/3)$: (a) find the mean; (b) find the variance and standard deviation; (c) sketch the shape of the probability distribution.

Q4. A binomial variable has mean 12 and variance $\frac{24}{5}$. (a) use $1 - p = \frac{\mathrm{Var}}{\mathrm{mean}}$ to find p; (b) hence find n.

Q5. For $X \sim \mathrm{Bi}(9, 2/3)$: (a) find the mean and variance; (b) describe the skew of the graph and explain why.

Q6. A fair coin is tossed 100 times and X is the number of heads. (a) state the distribution of X; (b) find the mean and standard deviation of X.

Unscaffolded

For **Q7–Q13** and **Q17–Q18**, find the mean, variance and standard deviation (exact values).

Q7. $X \sim \mathrm{Bi}(16, 1/4)$

Q8. $X \sim \mathrm{Bi}(50, 1/5)$

Q9. $X \sim \mathrm{Bi}(100, 1/2)$

Q10. $X \sim \mathrm{Bi}(24, 1/4)$

Q11. $X \sim \mathrm{Bi}(9, 2/3)$

Q12. $X \sim \mathrm{Bi}(40, 3/5)$

Q13. $X \sim \mathrm{Bi}(15, 1/5)$

Q14. For $X \sim \text{Bi}(5, 1/2)$, find $Pr(X=3)$ and state the mode.

Q15. For $X \sim \text{Bi}(4, 1/3)$, find $Pr(X=2)$.

Q16. Two binomial variables have the same n but $p=\frac{1}{5}$ and $p=\frac{4}{5}$. Compare the skew of their graphs and state how their variances compare.

Q17. $X \sim \text{Bi}(25, 4/5)$

Q18. $X \sim \text{Bi}(30, 1/6)$

Q19. A binomial variable has mean 8 and variance $\frac{8}{5}$. Find n and p.

Q20. A binomial variable has mean 20 and variance 16. Find n and p.

Q21. A binomial variable has mean 24 and standard deviation $\sqrt{6}$. Show that $p=\frac{3}{4}$ and find n.

Answers

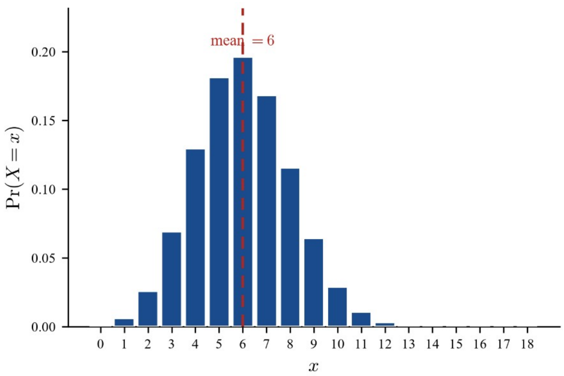
Q1. $q=\frac{3}{4}$; mean 3; variance $\frac{9}{4}$; sd $\frac{3}{2}$. **Q2.** Mean 10, variance 5; sd $\sqrt{5}$; symmetric. **Q3.** Mean 6; variance 4, sd 2; positively skewed. **Q4.** $p=\frac{3}{5}$; $n=20$. **Q5.** Mean 6, variance 2; negatively skewed (since $p=2/3>1/2$). **Q6.** $X\sim\mathrm{Bi}(100,1/2)$; mean 50, sd 5. **Q7.** Mean 4, variance 3, sd $\sqrt{3}$. **Q8.** Mean 10, variance 8, sd $2\sqrt{2}$. **Q9.** Mean 50, variance 25, sd 5. **Q10.** Mean 6, variance $\frac{9}{2}$, sd $\frac{3\sqrt{2}}{2}$. **Q11.** Mean 6, variance 2, sd $\sqrt{2}$. **Q12.** Mean 24, variance $\frac{48}{5}$, sd $\frac{4\sqrt{15}}{5}$. **Q13.** Mean 3, variance $\frac{12}{5}$, sd $\frac{2\sqrt{15}}{5}$. **Q14.** $Pr(X=3)=\frac{5}{16}$; modes 2 and 3 (both tallest). **Q15.** $Pr(X=2)=\frac{8}{27}$. **Q16.** $p=1/5$ is positively skewed, $p=4/5$ is negatively skewed (mirror images); the variances are **equal**, since $p(1-p)$ is the same for p and $1-p$. **Q17.** Mean 20, variance 4, sd 2. **Q18.** Mean 5, variance $\frac{25}{6}$, sd $\frac{5\sqrt{6}}{6}$. **Q19.** $p=\frac{4}{5}$, $n=10$. **Q20.** $p=\frac{1}{5}$, $n=100$. **Q21.** $p=\frac{3}{4}$, $n=32$.

Fully-worked solutions

Q1. $n=12$, $p=\frac{1}{4}$, $q=\frac{3}{4}$. Mean $=np=3$. Variance $=npq=12\times\frac{1}{4}\times\frac{3}{4}=\frac{9}{4}$. sd $=\sqrt{\frac{9}{4}}=\frac{3}{2}$.

Q2. Mean $=20\times\frac{1}{2}=10$. Variance $=20\times\frac{1}{2}\times\frac{1}{2}=5$. sd $=\sqrt{5}$. As $p=\frac{1}{2}$, the graph is **symmetric**.

Q3. Mean $=18\times\frac{1}{3}=6$. Variance $=18\times\frac{1}{3}\times\frac{2}{3}=4$, sd $=2$. Since $p=\frac{1}{3}<\frac{1}{2}$, the graph is positively skewed.

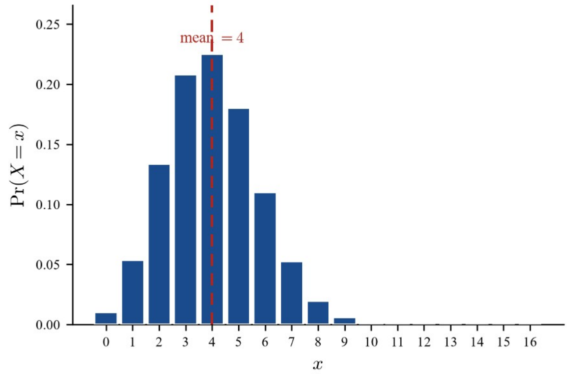


Q4. $1-p=\frac{\text{Var}}{\text{mean}}=\frac{24/5}{12}=\frac{2}{5}$, so $p=\frac{3}{5}$. Then $n=\frac{12}{3/5}=20$.

Q5. Mean $=9\times\frac{2}{3}=6$. Variance $=9\times\frac{2}{3}\times\frac{1}{3}=2$. Since $p=\frac{2}{3}>\frac{1}{2}$, the graph is **negatively skewed** (tail to the left).

Q6. Each toss is a Bernoulli trial with $p=\frac{1}{2}$, so $X\sim\mathrm{Bi}(100,1/2)$. Mean $=100\times\frac{1}{2}=50$; sd $=\sqrt{100\times1/2\times1/2}=\sqrt{25}=5$.

Q7. Mean $=16\times\frac{1}{4}=4$. Variance $=16\times\frac{1}{4}\times\frac{3}{4}=3$; sd $=\sqrt{3}$.

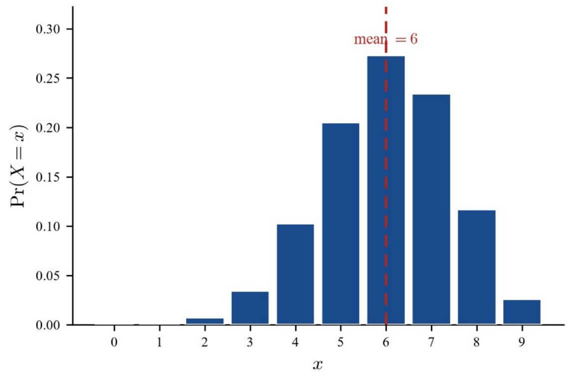


Q8. Mean $=50\times\frac{1}{5}=10$. Variance $=50\times\frac{1}{5}\times\frac{4}{5}=8$; sd $=\sqrt{8}=2\sqrt{2}$.

Q9. Mean $=100\times\frac{1}{2}=50$. Variance $=100\times\frac{1}{2}\times\frac{1}{2}=25$; sd $=5$.

Q10. Mean $=24\times\frac{1}{4}=6$. Variance $=24\times\frac{1}{4}\times\frac{3}{4}=\frac{9}{2}$; sd $=\sqrt{\frac{9}{2}}=\frac{3}{\sqrt{2}}=\frac{3\sqrt{2}}{2}$.

Q11. Mean $=9\times\frac{2}{3}=6$. Variance $=9\times\frac{2}{3}\times\frac{1}{3}=2$; sd $=\sqrt{2}$.



Q12. Mean $=40\times\frac{3}{5}=24$. Variance $=40\times\frac{3}{5}\times\frac{2}{5}=\frac{48}{5}$; sd $=\sqrt{\frac{48}{5}}=\frac{4\sqrt{3}}{\sqrt{5}}=\frac{4\sqrt{15}}{5}$.

Q13. Mean $=15\times\frac{1}{5}=3$. Variance $=15\times\frac{1}{5}\times\frac{4}{5}=\frac{12}{5}$; sd $=\sqrt{\frac{12}{5}}=\frac{2\sqrt{15}}{5}$.

Q14. $Pr(X=3)=\binom{5}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2=10\times\frac{1}{32}=\frac{5}{16}$. The graph is symmetric about 2.5, so both $x=2$ and $x=3$ are modes.

Q15. $Pr(X=2)=\binom{4}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^2=6\times\frac{1}{9}\times\frac{4}{9}=\frac{24}{81}=\frac{8}{27}$.

Q16. The $p=\frac{1}{5}$ graph is positively skewed and the $p=\frac{4}{5}$ graph is negatively skewed — one is the mirror image of the other. Their variances are **equal**, because $\mathrm{Var}=np(1-p)$ and $\frac{1}{5}\times\frac{4}{5}=\frac{4}{5}\times\frac{1}{5}$.

Q17. Mean $=25\times\frac{4}{5}=20$. Variance $=25\times\frac{4}{5}\times\frac{1}{5}=4$; sd $=2$.

Q18. Mean $=30\times\frac{1}{6}=5$. Variance $=30\times\frac{1}{6}\times\frac{5}{6}=\frac{25}{6}$; sd $=\sqrt{\frac{25}{6}}=\frac{5}{\sqrt{6}}=\frac{5\sqrt{6}}{6}$.

Q19. $1-p=\frac{8/5}{8}=\frac{1}{5}$, so $p=\frac{4}{5}$; $n=\frac{8}{4/5}=10$.

Q20. $1-p=\frac{16}{20}=\frac{4}{5}$, so $p=\frac{1}{5}$; $n=\frac{20}{1/5}=100$.

Q21. Variance $=\left(\sqrt{6}\right)^2=6$. Then $1-p=\frac{\text{Var}}{\text{mean}}=\frac{6}{24}=\frac{1}{4}$, so $p=\frac{3}{4}$. Hence $n=\frac{24}{3/4}=32$.

Theory

For a binomial random variable $X \sim \mathrm{Bi}(n, p)$ — the number of successes in n independent trials, each with success probability p — a very common question is:

> *How many trials are needed so that the probability of **at least one** success exceeds a given value c?*

The key is to work with the complement. Since $Pr(X \geq 1) = 1 - Pr(X = 0)$ and $Pr(X = 0) = \binom{n}{0} p^0 (1-p)^n = (1-p)^n$,

$$Pr(X \geq 1) = 1 - (1-p)^n.$$

Setting up the inequality. We require $Pr(X \geq 1) > c$, so

$$1 - (1-p)^n > c \Leftrightarrow (1-p)^n < 1 - c.$$

Solving for n with logarithms. Take $\log_e$ of both sides:

$$n \log_e(1-p) < \log_e(1-c).$$

Because $0 < 1 - p < 1$, the number $\log_e(1-p)$ is **negative**, so dividing by it **reverses** the inequality:

$$n > \frac{\log_e(1-c)}{\log_e(1-p)}.$$

Rounding. The right-hand side is usually not a whole number. Since n must be a positive integer and must be *greater* than this value, take the **next integer up** (round up). Always check the answer: n should satisfy the condition and $n - 1$ should fail it.

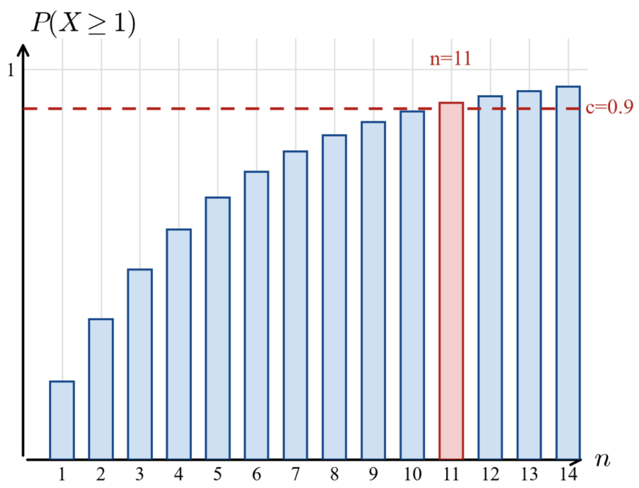
> **Note.** Watch the direction of the inequality carefully — dividing by the negative $\log_e(1-p)$ flips < to >. And n must be a whole number of trials, so round **up**, never down.

Worked examples

Example 1 — scaffolded

A spinner lands on red with probability $\frac{1}{5}$ on each spin. How many times must it be spun so that the probability of getting **at least one** red exceeds 0.9?

Working	Comment
Let X be the number of reds in n spins, $X \sim \mathrm{Bi}(n, 1/5)$.	Each spin is an independent trial with $p = 1/5$.
$Pr(X \geq 1) = 1 - (4/5)^n$. Require $1 - (4/5)^n > 0.9$.	Use the complement $Pr(X \geq 1) = 1 - Pr(X = 0)$.
$(4/5)^n < 0.1$	Rearrange.
$n \log_e(4/5) < \log_e(0.1)$, so $n > \frac{\log_e(0.1)}{\log_e(0.8)} \approx 10.32$.	Take $\log_e$; dividing by $\log_e(0.8) < 0$ **reverses** the inequality.
$n = 11$.	Round up to the next whole number.
Check: $1 - (0.8)^{11} \approx 0.914 > 0.9$ ✓, and $1 - (0.8)^{10} \approx 0.893 < 0.9$.	$n = 11$ works; $n = 10$ does not.



Example 2 — scaffolded

5 % of the items produced by a machine are defective. How many items must be inspected so that the probability of finding **at least one** defective is greater than 0.95?

Working	Comment
Let $X \sim \text{Bi}(n, 0.05)$ be the number of defectives in n items.	$p = 0.05$, so $1 - p = 0.95$.
Require $1 - (0.95)^n > 0.95$, i.e. $(0.95)^n < 0.05$.	Complement, then rearrange.
$n > \dfrac{\log_e(0.05)}{\log_e(0.95)} \approx 58.40$.	Divide by $\log_e(0.95) < 0 \rightarrow$ reverse the inequality.
$n = 59$.	Round up.
Check: $1 - (0.95)^{59} \approx 0.9515 > 0.95$ ✓; $1 - (0.95)^{58} \approx 0.9490 < 0.95$.	59 items are needed.

Example 3 — unscaffolded

An archer hits the target with probability 0.7 on each shot. How many shots are needed so that the probability of **at least one** hit exceeds 0.999?

Let $X \sim \text{Bi}(n, 0.7)$. Require $1 - (0.3)^n > 0.999$, so $(0.3)^n < 0.001$.

$$n > \frac{\log_e(0.001)}{\log_e(0.3)} \approx 5.74.$$

$\therefore n = 6$. (Check: $1 - (0.3)^6 \approx 0.99927 > 0.999$; $1 - (0.3)^5 \approx 0.99757 < 0.999$.)

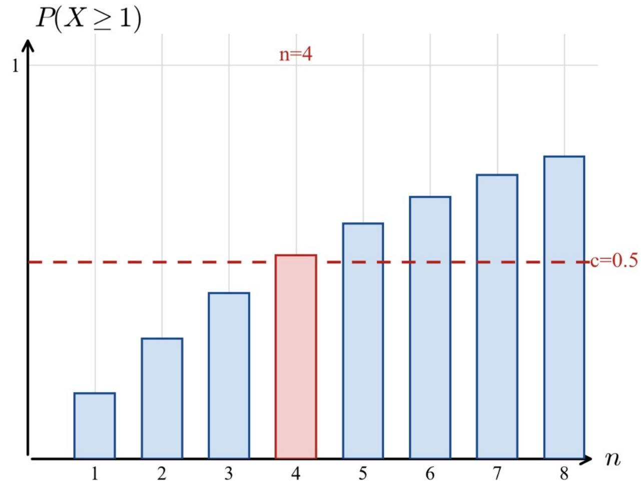
Example 4 — unscaffolded

A fair six-sided die is rolled repeatedly. How many rolls are needed so that the probability of getting **at least one** six is greater than 0.5?

Let $X \sim \text{Bi}(n, 1/6)$. Require $1 - (5/6)^n > 0.5$, so $(5/6)^n < 0.5$.

$$n > \frac{\log_e(0.5)}{\log_e(5/6)} \approx 3.80.$$

$\therefore n = 4$. (Check: $1 - (5/6)^4 \approx 0.518 > 0.5$; $1 - (5/6)^3 \approx 0.421 < 0.5$.)



Practice questions

Scaffolded

Q1. On each turn a game is won with probability $\frac{3}{10}$. (a) Write an expression for $Pr(\text{at least one win in } n \text{ turns})$; (b) form the inequality for at least one win with probability greater than 0.9; (c) solve for n using logarithms; (d) state the least number of turns.

Q2. 10% of chocolates in a box contain a nut. How many chocolates must be chosen (with replacement) so that $Pr(\text{at least one nut}) > 0.8$? Show the logarithm step and round appropriately.

Q3. A basketball player makes a free throw with probability $\frac{1}{4}$. Find the least number of attempts so that the probability of at least one success exceeds 0.95.

Q4. A biased coin shows heads with probability $\frac{2}{5}$. Find the smallest number of tosses so that the probability of at least one head is greater than 0.99.

Q5. A die is rolled repeatedly. Find the least number of rolls so that the probability of at least one six exceeds 0.9.

Q6. 15% of a large batch of seeds fail to germinate. How many seeds must be planted so that the probability that at least one fails to germinate is greater than 0.9?

Unscaffolded

In each case find the **smallest** sample size n satisfying the stated condition (round up to a whole number), showing the logarithm working.

Q7. A fair coin is tossed. $Pr(\text{at least one head}) > 0.999$.

Q8. $p = 0.2$; $Pr(X \geq 1) > 0.99$.

Q9. 8% of light globes are faulty; $Pr(\text{at least one faulty}) > 0.9$.

Q10. A component works with probability 0.6; $Pr(\text{at least one working}) > 0.9999$.

Q11. A dart hits the bullseye with probability 0.35; $Pr(\text{at least one bullseye}) > 0.95$.

Q12. 2% of tickets win a prize; $Pr(\text{at least one winning ticket}) > 0.5$.

Q13. $p = 0.45$; $Pr(X \geq 1) > 0.99$.

Q14. A spinner shows blue with probability $\frac{1}{4}$; $Pr(\text{at least one blue}) > 0.99$.

Q15. 12% of emails are spam; $Pr(\text{at least one spam}) > 0.8$.

Q16. 1% of items on a production line are defective; $Pr(\text{at least one defective}) > 0.9$.

Q17. 5% of a population carry a trait; $Pr(\text{at least one carrier in the sample}) > 0.9$.

Q18. $p = 0.3$; $Pr(X \geq 1) > 0.999$.

Q19. 18% of calls are dropped; $Pr(\text{at least one dropped call}) > 0.95$.

Q20. A die is rolled repeatedly; $Pr(\text{at least one six}) > 0.99$.

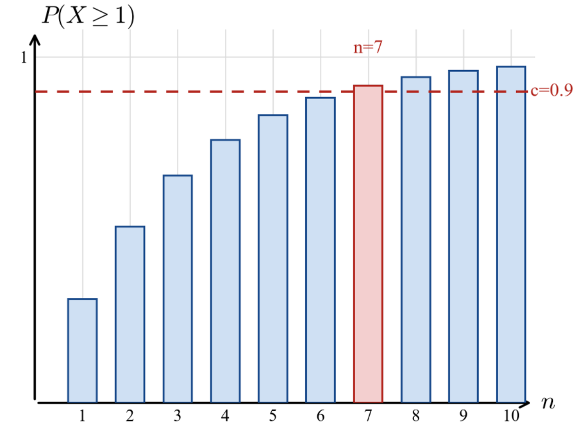
Q21. 22% of customers use a coupon; $Pr(\text{at least one coupon}) > 0.9$.

Answers

Q1. $n=7$. **Q2.** $n=16$. **Q3.** $n=11$. **Q4.** $n=10$. **Q5.** $n=13$. **Q6.** $n=15$. **Q7.** $n=10$. **Q8.** $n=21$. **Q9.** $n=28$. **Q10.** $n=11$. **Q11.** $n=7$. **Q12.** $n=35$. **Q13.** $n=8$. **Q14.** $n=17$. **Q15.** $n=13$. **Q16.** $n=230$. **Q17.** $n=45$. **Q18.** $n=20$. **Q19.** $n=16$. **Q20.** $n=26$. **Q21.** $n=10$.

Fully-worked solutions

Q1. (a) $Pr(\geq 1)=1-(7/10)^n$. (b) $1-(0.7)^n>0.9$, i.e. $(0.7)^n<0.1$. (c) $n>\dfrac{\log_e(0.1)}{\log_e(0.7)}\approx 6.46$. (d) $n=7$ (check $1-(0.7)^7\approx 0.918>0.9$; $1-(0.7)^6\approx 0.882$).



Q2. $(0.9)^n<0.2 \Rightarrow n>\dfrac{\log_e(0.2)}{\log_e(0.9)}\approx 15.28 \Rightarrow n=16$. $(1-(0.9)^{16}\approx 0.815>0.8.)$

Q3. $(3/4)^n<0.05 \Rightarrow n>\dfrac{\log_e(0.05)}{\log_e(0.75)}\approx 10.41 \Rightarrow n=11$. $(1-(0.75)^{11}\approx 0.958.)$

Q4. $(3/5)^n<0.01 \Rightarrow n>\dfrac{\log_e(0.01)}{\log_e(0.6)}\approx 9.02 \Rightarrow n=10$. $(1-(0.6)^{10}\approx 0.994.)$

Q5. $(5/6)^n<0.1 \Rightarrow n>\dfrac{\log_e(0.1)}{\log_e(5/6)}\approx 12.63 \Rightarrow n=13$. $(1-(5/6)^{13}\approx 0.907.)$

Q6. $(0.85)^n<0.1 \Rightarrow n>\dfrac{\log_e(0.1)}{\log_e(0.85)}\approx 14.17 \Rightarrow n=15$. $(1-(0.85)^{15}\approx 0.913.)$

Q7. $(1/2)^n<0.001 \Rightarrow n>\dfrac{\log_e(0.001)}{\log_e(0.5)}\approx 9.97 \Rightarrow n=10$. $(1-(0.5)^{10}\approx 0.99902.)$

Q8. $(0.8)^n<0.01 \Rightarrow n>\dfrac{\log_e(0.01)}{\log_e(0.8)}\approx 20.64 \Rightarrow n=21$. $(1-(0.8)^{21}\approx 0.9908.)$

Q9. $(0.92)^n<0.1 \Rightarrow n>\dfrac{\log_e(0.1)}{\log_e(0.92)}\approx 27.62 \Rightarrow n=28$. $(1-(0.92)^{28}\approx 0.903.)$

Q10. $(0.4)^n<0.0001 \Rightarrow n>\dfrac{\log_e(0.0001)}{\log_e(0.4)}\approx 10.05 \Rightarrow n=11$. $(1-(0.4)^{11}\approx 0.99996.)$

Q11. $(0.65)^n < 0.05 \Rightarrow n > \dfrac{\log_e(0.05)}{\log_e(0.65)} \approx 6.95 \Rightarrow n = 7$. ($1-(0.65)^7 \approx 0.951$.)

Q12. $(0.98)^n < 0.5 \Rightarrow n > \dfrac{\log_e(0.5)}{\log_e(0.98)} \approx 34.31 \Rightarrow n = 35$. ($1-(0.98)^{35} \approx 0.507$.)

Q13. $(0.55)^n < 0.01 \Rightarrow n > \dfrac{\log_e(0.01)}{\log_e(0.55)} \approx 7.70 \Rightarrow n = 8$. ($1-(0.55)^8 \approx 0.9916$.)

Q14. $(3/4)^n < 0.01 \Rightarrow n > \dfrac{\log_e(0.01)}{\log_e(0.75)} \approx 16.01 \Rightarrow n = 17$. ($1-(0.75)^{17} \approx 0.9925$.)

Q15. $(0.88)^n < 0.2 \Rightarrow n > \dfrac{\log_e(0.2)}{\log_e(0.88)} \approx 12.59 \Rightarrow n = 13$. ($1-(0.88)^{13} \approx 0.810$.)

Q16. $(0.99)^n < 0.1 \Rightarrow n > \dfrac{\log_e(0.1)}{\log_e(0.99)} \approx 229.11 \Rightarrow n = 230$. ($1-(0.99)^{230} \approx 0.9009$.)

Q17. $(0.95)^n < 0.1 \Rightarrow n > \dfrac{\log_e(0.1)}{\log_e(0.95)} \approx 44.89 \Rightarrow n = 45$. ($1-(0.95)^{45} \approx 0.9006$.)

Q18. $(0.7)^n < 0.001 \Rightarrow n > \dfrac{\log_e(0.001)}{\log_e(0.7)} \approx 19.37 \Rightarrow n = 20$. ($1-(0.7)^{20} \approx 0.99920$.)

Q19. $(0.82)^n < 0.05 \Rightarrow n > \dfrac{\log_e(0.05)}{\log_e(0.82)} \approx 15.10 \Rightarrow n = 16$. ($1-(0.82)^{16} \approx 0.958$.)

Q20. $(5/6)^n < 0.01 \Rightarrow n > \dfrac{\log_e(0.01)}{\log_e(5/6)} \approx 25.26 \Rightarrow n = 26$. ($1-(5/6)^{26} \approx 0.9913$.)

Q21. $(0.78)^n < 0.1 \Rightarrow n > \dfrac{\log_e(0.1)}{\log_e(0.78)} \approx 9.27 \Rightarrow n = 10$. ($1-(0.78)^{10} \approx 0.917$.)

Chapter 12 The binomial distribution — 12D Proving the mean and variance formulas

Extension. The study design requires the results $E(X)=np$ and $Var(X)=np(1-p)$, not their proofs. This section — including the binomial theorem it relies on — is included for students who want to see why the formulas hold, and is not assessed in the graded topic tests or practice examinations.

Theory

For a binomial random variable $X \sim \mathrm{Bi}(n, p)$ the probability function is

$$Pr(X=k)=\binom{n}{k}p^k q^{n-k}, k=0,1,2,\dots,n,$$

where $q=1-p$. This section proves the two standard results

$$E(X)=np \text{ and } Var(X)=np(1-p),$$

so that $\mathrm{sd}(X)=\sqrt{np(1-p)}$.

The binomial theorem. For any numbers a and b and positive integer n,

$$(a+b)^n=\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}=\binom{n}{0}b^n+\binom{n}{1}ab^{n-1}+\cdots+\binom{n}{n}a^n.$$

Each term arises by **choosing** which k of the n factors contribute an a — hence the $\binom{n}{k}$ ways. For example, $(a+b)^3=b^3+3ab^2+3a^2b+a^3$, with coefficients $\binom{3}{0},\binom{3}{1},\binom{3}{2},\binom{3}{3}=1,3,3,1$.

The probabilities sum to 1**.** By the binomial theorem, putting $a=p$ and $b=q$,

$$\sum_{k=0}^{n}\binom{n}{k}p^k q^{n-k}=(p+q)^n=1^n=1,$$

since $p+q=1$. So the distribution is valid.

A key identity. For $1\le k\le n$,

$$k\binom{n}{k}=n\binom{n-1}{k-1}.$$

This is because $k\dfrac{n!}{k!(n-k)!}=\dfrac{n!}{(k-1)!(n-k)!}=n\cdot\dfrac{(n-1)!}{(k-1)!(n-k)!}=n\dbinom{n-1}{k-1}$.

Proof that $E(X)=np$. The $k=0$ term contributes 0, so

$$E(X)=\sum_{k=0}^{n}k\binom{n}{k}p^k q^{n-k}=\sum_{k=1}^{n}n\binom{n-1}{k-1}p^k q^{n-k}.$$

Take out np and substitute $j=k-1$:

$$E(X)=np\sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1}q^{n-k}=np\sum_{j=0}^{n-1}\binom{n-1}{j}p^j q^{(n-1)-j}=np(p+q)^{n-1}=np.$$

Proof that $Var(X)=np(1-p)$. Use the companion identity, valid for $2\le k\le n$,

$$k(k-1)\binom{n}{k}=n(n-1)\binom{n-2}{k-2}.$$

The $k=0,1$ terms contribute 0, so

$$E(X(X-1))=\sum_{k=2}^{n} n(n-1)\binom{n-2}{k-2}p^k q^{n-k}=n(n-1)p^2\sum_{j=0}^{n-2}\binom{n-2}{j}p^j q^{(n-2)-j}=n(n-1)p^2.$$

Then, since $E(X^2)=E(X(X-1))+E(X)$,

$$E(X^2)=n(n-1)p^2+np,$$

and therefore

$$Var(X)=E(X^2)-[E(X)]^2=n(n-1)p^2+np-(np)^2=np-np^2=np(1-p).$$

Summary. For $X\sim\mathrm{Bi}(n,p)$: $E(X)=np$, $Var(X)=np(1-p)=npq$, $\mathrm{sd}(X)=\sqrt{np(1-p)}$.

Worked examples

Example 1 — scaffolded

For $X\sim\mathrm{Bi}(3,p)$, verify by direct summation that $E(X)=3p$.

Working	Comment
$Pr(X=0)=q^3, Pr(X=1)=3pq^2, Pr(X=2)=3p^2q,$	The four terms of $\binom{3}{k}p^k q^{3-k}$.
$E(X)=0\cdot q^3+1\cdot 3pq^2+2\cdot 3p^2q+3\cdot p^3$	$E(X)=\sum k\,Pr(X=k)$; the $k=0$ term is 0.
$=3pq^2+6p^2q+3p^3=3p(q^2+2pq+p^2)$	Factor out $3p$.
$=3p(q+p)^2=3p\cdot 1^2$	$q^2+2pq+p^2=(q+p)^2$ and $p+q=1$.
$\therefore E(X)=3p$	Agrees with np for $n=3$.

Example 2 — scaffolded

Verify the identity $k\binom{n}{k}=n\binom{n-1}{k-1}$ for $n=5, k=2$, and state the general reason.

Working	Comment
LHS $=2\binom{5}{2}=2\times 10=20$	$\binom{5}{2}=10$.
RHS $=5\binom{4}{1}=5\times 4=20$	$\binom{4}{1}=4$.
LHS = RHS $=20$	The identity holds here.
In general	Cancel one factor of k; then split off n.
$k\dfrac{n!}{k!(n-k)!}=\dfrac{n!}{(k-1)!(n-k)!}=n\dfrac{(n-1)!}{(k-1)!(n-k)!}$	
$\therefore k\binom{n}{k}=n\binom{n-1}{k-1}$	True for every $1\le k\le n$.

Example 3 — unscaffolded

Prove that $E(X)=np$ for $X\sim\mathrm{Bi}(n,p)$.

$E(X)=\sum_{k=0}^{n} k\binom{n}{k} p^k q^{n-k}=\sum_{k=1}^{n} n\binom{n-1}{k-1} p^k q^{n-k}$, using $k\binom{n}{k}=n\binom{n-1}{k-1}$ (the $k=0$ term is 0).

Taking out $n\,p$ and letting $j=k-1$:

$$E(X)=n\,p\sum_{j=0}^{n-1}\binom{n-1}{j} p^j q^{(n-1)-j}=n\,p\,(p+q)^{n-1}=n\,p\;\cdot 1.$$

$\therefore E(X)=n\,p.$

Example 4 — unscaffolded

Prove that $Var(X)=n\,p(1-p)$ for $X\sim\mathrm{Bi}(n,p)$.

Using $k(k-1)\binom{n}{k}=n(n-1)\binom{n-2}{k-2}$ (the $k=0,1$ terms are 0):

$$E\big(X(X-1)\big)=\sum_{k=2}^{n} n(n-1)\binom{n-2}{k-2} p^k q^{n-k}=n(n-1)\,p^2(p+q)^{n-2}=n(n-1)\,p^2.$$

Since $E(X^2)=E\big(X(X-1)\big)+E(X)=n(n-1)\,p^2+n\,p$,

$$Var(X)=E(X^2)-[E(X)]^2=n(n-1)\,p^2+n\,p-n^2p^2=n\,p-n\,p^2.$$

$\therefore Var(X)=n\,p(1-p).$

Practice questions

Scaffolded

Q1. For the identity $k\binom{n}{k}=n\binom{n-1}{k-1}$: (a) verify it for $n=6,k=3$; (b) verify it for $n=7,k=1$; (c) explain why the $k=0$ term may be dropped from the sum for $E(X)$.

Q2. For $X\sim\mathrm{Bi}(4,p)$: (a) write $\sum_{k=0}^{4}\binom{4}{k}p^k q^{4-k}$ in expanded form; (b) recognise this as a power of $(p+q)$; (c) hence show the probabilities sum to 1.

Q3. For $X\sim\mathrm{Bi}(2,p)$: (a) write down $Pr(X=0), Pr(X=1), Pr(X=2)$; (b) compute $E(X)$ by direct summation; (c) confirm that $E(X)=2\,p$.

Q4. For $X\sim\mathrm{Bi}(4,p)$, compute $E\big(X(X-1)\big)$ by direct summation and show that it equals $12\,p^2$.

Q5. Given $E\big(X(X-1)\big)=n(n-1)\,p^2$ and $E(X)=n\,p$: (a) write $E(X^2)$ in terms of n and p; (b) hence show $Var(X)=n\,p(1-p)$.

Q6. Verify the identity $k(k-1)\binom{n}{k}=n(n-1)\binom{n-2}{k-2}$ for $n=5,k=3$.

Unscaffolded

Q7. Verify that $k\binom{n}{k}=n\binom{n-1}{k-1}$ for $n=8,k=4$.

Q8. For $X\sim\mathrm{Bi}(3,p)$, compute $E(X^2)$ by direct summation and hence find $Var(X)$.

Q9. Show that $\sum_{k=0}^{5}\binom{5}{k}p^k q^{5-k}=1$, stating the theorem used.

Q10. Prove that $E(X)=np$ for $X\sim\text{Bi}(n,p)$, using the identity $k\binom{n}{k}=n\binom{n-1}{k-1}$.

Q11. Prove that $Var(X)=np(1-p)$ for $X\sim\text{Bi}(n,p)$.

Q12. For $X\sim\text{Bi}(4,p)$, verify by direct summation that $E(X)=4p$.

Q13. Prove the identity $k(k-1)\binom{n}{k}=n(n-1)\binom{n-2}{k-2}$ algebraically.

Q14. Using $E(X)=np$ and $Var(X)=np(1-p)$, find the mean, variance and standard deviation of $X\sim\text{Bi}\left(6,\frac{1}{4}\right)$, giving exact values.

Q15. Find the mean, variance and standard deviation of $X\sim\text{Bi}\left(10,\frac{3}{10}\right)$, giving exact values.

Q16. Find the mean, variance and standard deviation of $X\sim\text{Bi}\left(8,\frac{1}{4}\right)$, giving exact values.

Q17. Show that for $X\sim\text{Bi}(n,p)$, $Var(X)=E(X)(1-p)$, and hence explain why $Var(X)<E(X)$.

Q18. A binomial random variable has $E(X)=4$ and $Var(X)=\frac{12}{5}$. Find the values of n and p.

Q19. A binomial random variable has mean 6 and variance $\frac{9}{2}$. Find the values of n and p.

Q20. Show that $E(X^2)=n(n-1)p^2+np$ for $X\sim\text{Bi}(n,p)$, and hence find $E(X^2)$ for $X\sim\text{Bi}\left(5,\frac{2}{5}\right)$.

Q21. For $X\sim\text{Bi}(n,p)$ with n fixed, show that $\text{sd}(X)=\sqrt{np(1-p)}$ is greatest when $p=\frac{1}{2}$.

Answers

Q1. (a) $3\binom{6}{3}=60=6\binom{5}{2}$; (b) $1\binom{7}{1}=7=7\binom{6}{0}$; (c) the factor $k=0$ makes the term 0. **Q2.** $(p+q)^4=1$, so the sum is 1. **Q3.** (a) $q^2, 2pq, p^2$; (b),(c) $E(X)=2pq+2p^2=2p(q+p)=2p$. **Q4.** $E(X(X-1))=12p^2$. **Q5.** (a) $E(X^2)=n(n-1)p^2+np$; (b) $Var(X)=np(1-p)$. **Q6.** $6\binom{5}{3}=60=20\binom{3}{1}$; both equal 60. **Q7.** $4\binom{8}{4}=280=8\binom{7}{3}$. **Q8.** $E(X^2)=6p^2+3p$; $Var(X)=3p(1-p)$. **Q9.** $(p+q)^5=1$ (binomial theorem). **Q10.** Proof: $E(X)=np$. **Q11.** Proof: $Var(X)=np(1-p)$. **Q12.** $E(X)=4p$. **Q13.** Proof of the identity. **Q14.** $E(X)=\frac{3}{2}, Var(X)=\frac{9}{8}, \text{sd}(X)=\frac{3\sqrt{2}}{4}$. **Q15.** $E(X)=3, Var(X)=\frac{21}{10}, \text{sd}(X)=\frac{\sqrt{210}}{10}$. **Q16.** $E(X)=2, Var(X)=\frac{3}{2}, \text{sd}(X)=\frac{\sqrt{6}}{2}$. **Q17.** $Var(X)=np(1-p)=np\cdot(1-p)=E(X)(1-p)$; since $0<1-p<1$, $Var(X)<E(X)$. **Q18.** $n=10, p=\frac{2}{5}$. **Q19.** $n=24, p=\frac{1}{4}$. **Q20.** $E(X^2)=n(n-1)p^2+np$; for $\text{Bi}(5, 2/5)$, $E(X^2)=\frac{26}{5}$. **Q21.** $p(1-p)$ is greatest at $p=\frac{1}{2}$, so $\text{sd}(X)$ is greatest there.

Fully-worked solutions

Q1. (a) LHS $=3\binom{6}{3}=3\times 20=60$; RHS $=6\binom{5}{2}=6\times 10=60$. Equal. (b) LHS $=1\binom{7}{1}=7$; RHS $=7\binom{6}{0}=7\times 1=7$. Equal. (c) In $E(X)=\sum_{k=0}^{n} k\,Pr(X=k)$ the $k=0$ term is $0\cdot Pr(X=0)=0$, so summation from $k=1$ gives the same value and lets us apply $k\binom{n}{k}=n\binom{n-1}{k-1}$.

Q2. (a) $\binom{4}{0}q^4+\binom{4}{1}pq^3+\binom{4}{2}p^2q^2+\binom{4}{3}p^3q+\binom{4}{4}p^4=q^4+4pq^3+6p^2q^2+4p^3q+p^4$. (b) This is the expansion of $(p+q)^4$. (c) $(p+q)^4=1^4=1$ since $p+q=1$, so the probabilities sum to 1.

Q3. (a) $Pr(X=0)=q^2, Pr(X=1)=2pq, Pr(X=2)=p^2$. (b) $E(X)=0\cdot q^2+1\cdot 2pq+2\cdot p^2=2pq+2p^2$. (c) $=2p(q+p)=2p\cdot 1=2p$.

Q4. For $\text{Bi}(4,p)$, $E(X(X-1))=\sum_{k=0}^{4} k(k-1)\binom{4}{k}p^k q^{4-k}$. Only $k=2,3,4$ contribute:
$2\cdot 1\cdot 6p^2q^2+3\cdot 2\cdot 4p^3q+4\cdot 3\cdot p^4=12p^2q^2+24p^3q+12p^4=12p^2(q^2+2pq+p^2)=12p^2(q+p)^2=12p^2$.

Q5. (a) $E(X^2)=E(X(X-1))+E(X)=n(n-1)p^2+np$. (b) $Var(X)=E(X^2)-[E(X)]^2=n(n-1)p^2+np-(np)^2=n^2p^2-np^2+np-n^2p^2=np-np^2=np(1-p)$.

Q6. LHS $=3\cdot 2\cdot\binom{5}{3}=6\times 10=60$; RHS $=5\cdot 4\cdot\binom{3}{1}=20\times 3=60$. Equal.

Q7. LHS $=4\binom{8}{4}=4\times 70=280$; RHS $=8\binom{7}{3}=8\times 35=280$. Equal.

Q8. For $\text{Bi}(3,p)$: $E(X(X-1))=6p^2$ (from Example/Q4 pattern with $n=3$: $2\cdot 3p^2q+6\cdot p^3=6p^2(q+p)=6p^2$). So $E(X^2)=6p^2+E(X)=6p^2+3p$. Then $Var(X)=6p^2+3p-(3p)^2=3p-3p^2=3p(1-p)$.

Q9. By the binomial theorem $\sum_{k=0}^{5}\binom{5}{k}p^k q^{5-k}=(p+q)^5=1^5=1$.

Q10. $E(X)=\sum_{k=0}^{n}k\binom{n}{k}p^k q^{n-k}=\sum_{k=1}^{n}n\binom{n-1}{k-1}p^k q^{n-k}$ (identity; $k=0$ term 0)
$=np\sum_{j=0}^{n-1}\binom{n-1}{j}p^j q^{(n-1)-j}=np(p+q)^{n-1}=np$.

Q11. Using $k(k-1)\binom{n}{k}=n(n-1)\binom{n-2}{k-2}$: $E(X(X-1))=n(n-1)p^2(p+q)^{n-2}=n(n-1)p^2$. Then
$E(X^2)=n(n-1)p^2+np$ and $Var(X)=E(X^2)-(np)^2=np-np^2=np(1-p)$.

Q12. For $\text{Bi}(4,p)$:
$E(X)=1\cdot 4pq^3+2\cdot 6p^2q^2+3\cdot 4p^3q+4\cdot p^4=4pq^3+12p^2q^2+12p^3q+4p^4=4p(q^3+3pq^2+3p^2q+p^3)=4p(q$
.

Q13. $k(k-1)\dfrac{n!}{k!(n-k)!}=\dfrac{n!}{(k-2)!(n-k)!}=n(n-1)\dfrac{(n-2)!}{(k-2)!(n-k)!}=n(n-1)\binom{n-2}{k-2}$, cancelling $k(k-1)$ against $k!$ and splitting off $n(n-1)$.

Q14. $E(X)=6\cdot\frac{1}{4}=\frac{3}{2}$; $Var(X)=6\cdot\frac{1}{4}\cdot\frac{3}{4}=\frac{18}{16}=\frac{9}{8}$; $\text{sd}(X)=\sqrt{\frac{9}{8}}=\frac{3}{2\sqrt{2}}=\frac{3\sqrt{2}}{4}$.

Q15. $E(X)=10\cdot\frac{3}{10}=3$; $Var(X)=10\cdot\frac{3}{10}\cdot\frac{7}{10}=\frac{21}{10}$; $\text{sd}(X)=\sqrt{\frac{21}{10}}=\frac{\sqrt{210}}{10}$.

Q16. $E(X)=8\cdot\frac{1}{4}=2$; $Var(X)=8\cdot\frac{1}{4}\cdot\frac{3}{4}=\frac{3}{2}$; $\text{sd}(X)=\sqrt{\frac{3}{2}}=\frac{\sqrt{6}}{2}$.

Q17. $Var(X)=np(1-p)=(np)(1-p)=E(X)(1-p)$. Since $0<p<1$ gives $0<1-p<1$, we have $Var(X)=E(X)(1-p)<E(X)$.

Q18. $np=4$ and $np(1-p)=\frac{12}{5}$. Dividing: $1-p=\frac{12/5}{4}=\frac{3}{5}$, so $p=\frac{2}{5}$, and $n=\frac{4}{p}=10$. $\therefore n=10,\ p=\frac{2}{5}$.

Q19. $np=6$ and $np(1-p)=\frac{9}{2}$. Dividing: $1-p=\frac{9/2}{6}=\frac{3}{4}$, so $p=\frac{1}{4}$, and $n=\frac{6}{p}=24$. $\therefore n=24,\ p=\frac{1}{4}$.

Q20. $E(X^2)=E(X(X-1))+E(X)=n(n-1)p^2+np$. For $\text{Bi}(5,2/5)$:
$E(X^2)=5\cdot 4\cdot\frac{4}{25}+5\cdot\frac{2}{5}=\frac{80}{25}+2=\frac{16}{5}+2=\frac{26}{5}$.

Q21. $\text{sd}(X)=\sqrt{np(1-p)}$; with n fixed this is greatest when $p(1-p)$ is greatest. $p(1-p)=\frac{1}{4}-\left(p-\frac{1}{2}\right)^2$, which is maximised at $p=\frac{1}{2}$ (value $\frac{1}{4}$). $\therefore$ the standard deviation is greatest when $p=\frac{1}{2}$.