Freebooks.pk

Free Notes • Past Papers • Guides for Students

ICS Part-I Statistics — Punjab Board

CHAPTER 9

Binomial & Hypergeometric Distribution

Statistics • 1st Year • ICS • Punjab Board • Free PDF Notes

Chapter 9: Binomial and Hypergeometric Probability Distribution

Chapter 8 built the general tools for describing any probability distribution -- the PMF, PDF, expectation, and variance. This chapter applies those tools to two of the most important and widely used discrete probability distributions in statistics: the binomial distribution, which models the number of successes in a fixed number of independent trials (like counting heads in repeated coin tosses, or defective items when sampling with replacement), and the hypergeometric distribution, which models the same kind of counting problem but when sampling is done without replacement from a finite population, making successive trials dependent rather than independent.

Both distributions answer the same basic question -- 'how many successes out of n trials?' -- but they apply to different real-world sampling situations, and recognizing which one fits a given problem (replacement vs. no replacement, fixed vs. changing probability of success) is one of the most practical skills this chapter builds.

Learning Objectives

- Define a Bernoulli trial and list the four properties of a binomial experiment
- Apply the binomial probability formula P(X=x) = C(n,x) p^x q^(n-x) to compute exact, cumulative, and range probabilities
- Identify whether a binomial distribution is symmetric, positively skewed, or negatively skewed based on p
- Calculate expected (binomial) frequencies for a fixed number of repeated experiments
- Derive and apply the mean (np) and variance (npq) of the binomial distribution
- Define a hypergeometric experiment and list its four defining properties
- Apply the hypergeometric probability formula and calculate its mean and variance

- Distinguish when to use the binomial distribution versus the hypergeometric distribution for a given sampling problem

Key Concepts

9.1 Introduction: Bernoulli Trials and the Binomial Experiment

In experiments like tossing a coin or repeatedly drawing a card, each individual repetition is called a trial, and the result of each trial is classified as either a success or a failure. The probability of success is denoted p and the probability of failure is denoted q, where q = 1-p (so p+q=1). If an experiment is repeated n times, the number of successes is denoted x and the number of failures is n-x.

A trial with exactly two possible outcomes -- success or failure -- is called a Bernoulli trial; for each Bernoulli trial the probability of success stays the same, and successive trials are independent. An experiment made up of repeated Bernoulli trials, where the probability of success stays constant from trial to trial, is called a binomial experiment. A binomial experiment has four defining properties: (i) each trial results in an outcome classified as success or failure; (ii) the probability of success remains constant from one trial to the next; (iii) the repeated trials are independent of each other; (iv) the experiment is repeated a fixed number of times.

9.2 Binomial Probability Distribution

For n independent trials, each with probability of success p and probability of failure q (q+p=1), the probability of exactly x successes is P(X=x) = C(n,x).p^x.q^(n-x), for x=0,1,2,...,n. The random variable X is called the binomial variable, its distribution is the binomial distribution, and n and p are the two parameters of the distribution, often written b(x,n,p). These probabilities are exactly the individual terms produced by expanding the binomial expression (q+p)^n -- which is why the distribution is named 'binomial' and why the coefficients C(n,x) are called binomial coefficients.

The shape of the binomial distribution depends entirely on p: if p=q=1/2, the distribution is perfectly symmetrical. If p is not equal to q, the distribution is skewed -- specifically, if p is greater than 1/2 the distribution is negatively skewed (tail toward lower values), and if p is less than 1/2 the distribution is positively skewed (tail toward higher values).

9.2.1 Binomial Frequency Distribution

When a binomial probability distribution is multiplied by the number of experiments N (repetitions of the whole n-trial experiment), the result is called the binomial frequency distribution. The expected frequency of exactly x successes across N experiments is f(X=x) = N.P(X=x) = N.C(n,x).q^(n-x).p^x -- this is simply the theoretical probability scaled up to a real count, useful for predicting how many out of N groups (families, batches, samples) will show a particular number of successes.

9.2.2 Mean and Variance of the Binomial Distribution

Through algebraic derivation starting from E(X) = sum of x.P(x) and expanding using the binomial theorem, the mean of the binomial distribution simplifies elegantly to E(X) = np -- the number of trials multiplied by the probability of success on each trial, which matches the intuitive idea that if you flip a coin n times with probability p of heads, you'd 'expect' about np heads on average.

Following the same style of derivation for E(X^2), and using the computing formula Var(X) = E(X^2) - [E(X)]^2, the variance of the binomial distribution simplifies to Var(X) = npq, and consequently the standard deviation is sigma = square-root(npq). These two compact formulas, mean = np and variance = npq, are among the most frequently used results in this chapter and let you find the mean and spread of a binomial distribution instantly, without building a full probability table.

9.3 Hypergeometric Distribution

When successive trials are conducted without replacement, they are no longer independent -- the probability of success changes from one trial to the next depending on what was already drawn. An experiment where a random sample is chosen without replacement from a finite population is called a hypergeometric experiment. It has four defining properties: (i) the experiment is repeated a fixed number of times; (ii) the successive trials are dependent; (iii) the probability of success varies from trial to trial (it is not fixed); (iv) each outcome is still classified as success or failure. The random variable representing the number of successes is called the hypergeometric variable, and its distribution is the hypergeometric distribution.

Suppose there are N total items, of which k are classified as successes and (N-k) as failures, and n items are selected at random without replacement (n <= N). The probability of getting exactly x successes (and n-x failures) is P(X=x) = [C(k,x).C(N-k,n-x)] / C(N,n), for x=0,1,2,...,n. The hypergeometric distribution has three parameters: n, k, and N. Its mean is nk/N, and its variance is (nk/N).(1 - k/N).((N-n)/(N-1)) -- notice the extra factor (N-n)/(N-1), called the finite population correction factor, which distinguishes it from the simpler binomial variance and accounts for the fact that sampling is without replacement from a limited population.

Important Definitions

What is a Bernoulli trial?: A trial with exactly two possible outcomes -- success and failure -- where the probability of success stays constant and successive trials are independent.

What is a binomial experiment?: An experiment made up of a fixed number of independent Bernoulli trials, each with the same probability of success p.

What is the binomial probability formula?: P(X=x) = C(n,x).p^x.q^(n-x), giving the probability of exactly x successes in n independent trials.

What are the two parameters of the binomial distribution?: n (the number of trials) and p (the probability of success on each trial).

What is the mean and variance of the binomial distribution?: Mean = np; Variance = npq (where q = 1-p); Standard deviation = square-root(npq).

What is a hypergeometric experiment?: An experiment in which a random sample is drawn without replacement from a finite population, making successive trials dependent and the probability of success variable.

What is the hypergeometric probability formula?: P(X=x) = [C(k,x).C(N-k,n-x)] / C(N,n), where N is the population size, k is the number of successes in the population, and n is the sample size.

What is the finite population correction factor?: The term (N-n)/(N-1) appearing in the hypergeometric variance formula, which adjusts for sampling without replacement from a finite population of size N.

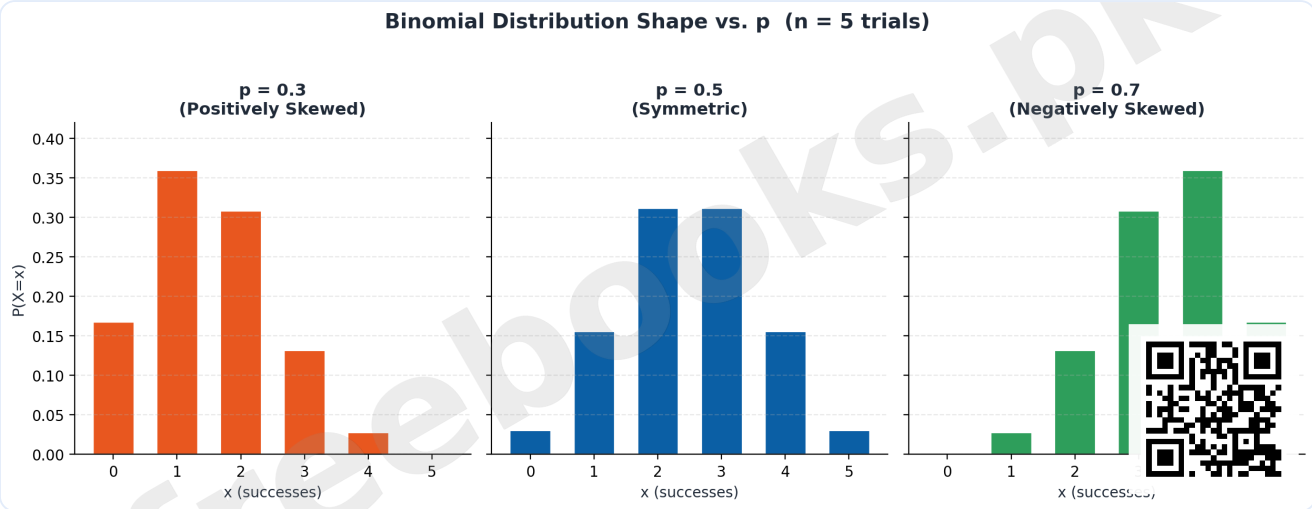
How do you decide whether to use the binomial or hypergeometric distribution?: Use the binomial distribution when sampling is with replacement (or from an effectively infinite population) so trials are independent and p is constant; use the hypergeometric distribution when sampling is without replacement from a small, finite population so trials are dependent and the probability of success changes.

Key Facts and Relations

Topic	Key Fact / Relation
Binomial probability	P(X=x) = C(n,x) p^x q^(n-x), x = 0,1,...,n
Binomial frequency distribution	f(X=x) = N.P(X=x) = N.C(n,x) q^(n-x) p^x
Binomial mean	E(X) = np
Binomial variance	Var(X) = npq
Binomial standard deviation	sigma = sqrt(npq)
Binomial distribution shape	p=q=0.5: symmetric | p>0.5: negatively skewed | p<0.5: positively skewed
Hypergeometric probability	P(X=x) = [C(k,x).C(N-k,n-x)] / C(N,n)
Hypergeometric mean	Mean = nk/N
Hypergeometric variance	Var(X) = (nk/N)(1 - k/N)((N-n)/(N-1))

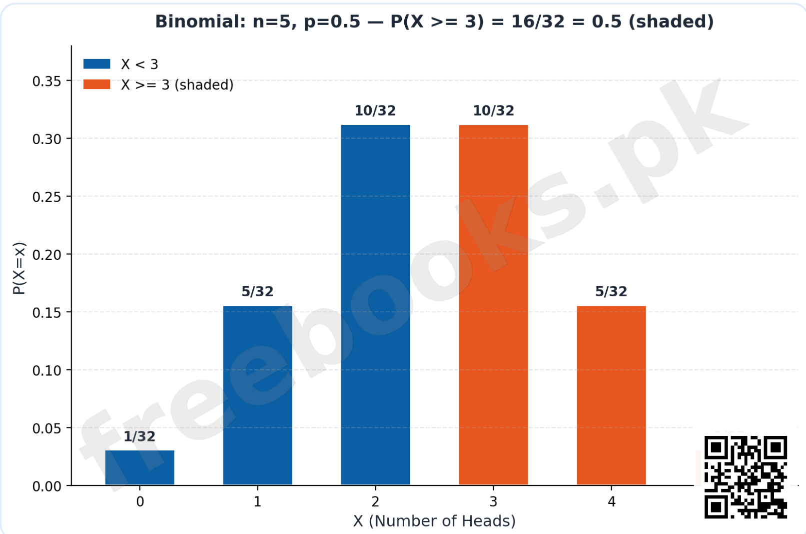
Diagrams

Binomial Distribution Shape vs. p (n=5): Three side-by-side binomial probability bar charts for n=5 with p=0.3, p=0.5, and p=0.7, illustrating positively skewed, symmetric, and negatively skewed binomial distributions respectively



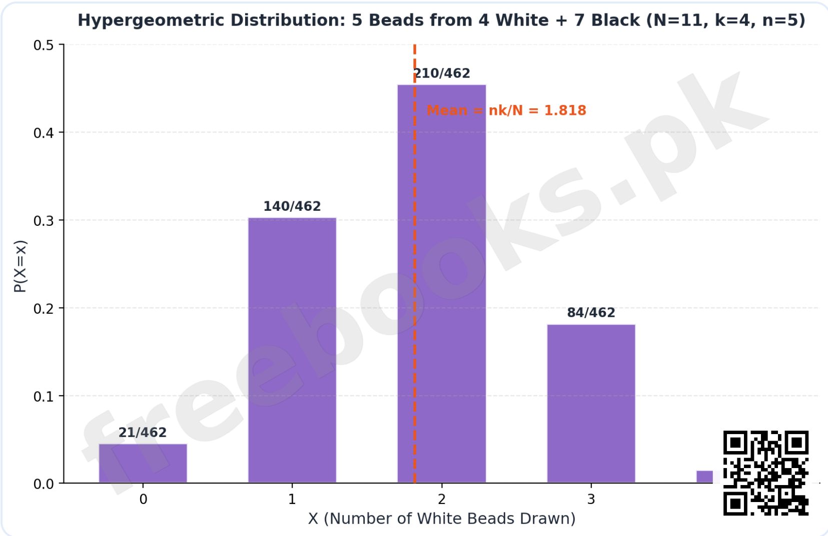
Binomial Distribution Shape vs. p (n=5)

Binomial Probability: At Least 3 Heads in 5 Coin Tosses: A binomial bar chart for n=5, p=0.5 (5 fair coin tosses) with the bars for X=3, X=4, X=5 shaded to show P(X>=3) = 16/32 = 0.5, matching Example 9.2 from the textbook



Binomial Probability: At Least 3 Heads in 5 Coin Tosses

Hypergeometric Probability Distribution: White Beads Example: A probability bar chart for the hypergeometric distribution of white beads drawn (5 beads from a bowl of 4 white and 7 black, N=11, k=4, n=5), with the mean (1.818) marked, matching Example 9.9 from the textbook



Hypergeometric Probability Distribution: White Beads Example

Short Questions & Answers

List the four properties of a binomial experiment.

Each trial results in success or failure; the probability of success stays constant across trials; the trials are independent; the experiment is repeated a fixed number of times.

Why is the binomial distribution symmetric when p = 0.5?

Because when p equals q (both 0.5), the probability of getting x successes equals the probability of getting (n-x) successes, making the distribution perfectly mirror-symmetric around its center.

State the mean and variance of a binomial distribution with n=20 trials and p=0.4.

Mean = np = 20(0.4) = 8; Variance = npq = 20(0.4)(0.6) = 4.8.

Why are successive trials dependent in a hypergeometric experiment but independent in a binomial experiment?

Because a hypergeometric experiment samples without replacement, so removing an item changes the composition of what remains and shifts the probability of success on the next draw; a binomial experiment samples with replacement (or from an effectively infinite population), so the probability of success never changes.

What does the finite population correction factor (N-n)/(N-1) account for in the hypergeometric variance?
It adjusts the variance downward to reflect that sampling without replacement from a finite population produces less variability than sampling with replacement would.

How many parameters does the binomial distribution have, and what are they? How about the hypergeometric?
Binomial: two parameters, n and p. Hypergeometric: three parameters, n, k, and N.

Long Questions & Answers

Explain the binomial probability distribution: define a binomial experiment, state its formula, and explain how its mean and variance are used, including how the shape of the distribution depends on p.

The binomial distribution is one of the most widely applied discrete probability distributions in all of statistics, precisely because it models an extremely common real-world situation: repeating the same simple yes-or-no experiment a fixed number of times under conditions that stay constant throughout. To understand why the binomial distribution takes the particular mathematical form it does, it helps to start from its most basic building block, the Bernoulli trial, which is simply any single trial with exactly two possible outcomes, conventionally labeled success and failure, where the probability of success, denoted p, stays fixed and where each trial has no influence whatsoever on any other trial. A binomial experiment is then nothing more than a fixed number, n, of these identical, independent Bernoulli trials strung together, and for a sequence of trials to genuinely qualify as a binomial experiment, four conditions must all hold simultaneously: every individual trial must produce an outcome that can be cleanly classified as either a success or a failure; the probability of success, p, must remain exactly the same from the very first trial all the way through to the very last trial; every trial must be completely independent of every other trial, meaning the outcome of one trial provides absolutely no information about the outcome of any other; and the total number of trials, n, must be fixed and known in advance before the experiment begins, rather than being allowed to vary. Given these conditions, the probability of observing exactly x successes across all n trials is given by the binomial probability formula, P(X=x) equals C(n,x) multiplied by p raised to the power x, multiplied by q raised to the power (n-x), where q simply represents the probability of failure and is always calculated as 1 minus p. The combinatorial term C(n,x) appearing at the front of this formula exists for a very specific and important reason: it counts the number of distinct ways that exactly x successes could be arranged among the n total trials, since a run of x successes could occur in many different orders across the sequence of trials, and every one of those equally likely orderings needs to be included in the total probability. This entire formula, and indeed the very name 'binomial' distribution itself, traces directly back to the algebraic expansion of the binomial expression (q+p) raised to the power n, since expanding this expression using the binomial theorem produces exactly these same terms, C(n,x).q^(n-x).p^x, as its individual components -- which is also precisely why the combinatorial coefficients C(n,x) are traditionally referred to as binomial coefficients. Beyond simply calculating the probability of one specific exact number of successes, the binomial formula also directly enables answering more complex 'at least,' 'at most,' or 'between' style questions, simply by calculating the individual probabilities for every relevant value of x and adding them together, since these composite events are always mutually exclusive when x is restricted to whole numbers. Once a specific binomial distribution has been fully defined by choosing its two parameters, n and p, two remarkably clean summary formulas become available without needing to build out an entire probability table from scratch. Through a somewhat lengthy but ultimately elegant algebraic derivation that expands E(X) as a weighted sum and repeatedly applies the underlying binomial theorem, the mean of any binomial distribution simplifies down to the strikingly simple result E(X) equals n times p -- a result that lines up perfectly with basic intuition, since if a fair coin with a 50% chance of heads is flipped 10 times, we would naturally expect around 10 times 0.5, or 5 heads, on average across many repetitions of that same 10-flip experiment. Following essentially the same style of derivation for E(X-squared), and then substituting the results into the standard computing formula Var(X) equals E(X-squared) minus [E(X)] squared, the variance of the binomial distribution likewise reduces down to the compact and equally memorable result Var(X) equals n.p.q, with the standard deviation then simply following as the square root of n.p.q. Beyond their obvious practical usefulness for quickly summarizing a binomial distribution's center and spread without laboriously constructing a full probability table, the parameter p specifically also governs the overall visual shape of the distribution whenever it gets plotted out as a bar chart or histogram. Whenever p happens to equal q, which necessarily means both must equal exactly 0.5, the resulting binomial distribution comes out perfectly symmetrical, since under these particular conditions the probability of observing x successes always exactly equals the probability of observing (n minus x) successes, causing the entire distribution to mirror itself precisely around its own central peak. Whenever p and q are not equal to each other, however, the distribution instead becomes skewed in one particular direction or the other, and the specific direction of that skew depends entirely and predictably on whether p happens to be larger or smaller than 0.5: specifically, whenever p is greater than 0.5, meaning success is inherently the comparatively more likely outcome on any single individual trial, the resulting distribution becomes negatively skewed, with its longer statistical tail stretching out toward the lower, smaller possible values of x; conversely, whenever p is less than 0.5, meaning success is instead the comparatively less likely outcome on any individual trial, the distribution becomes positively skewed instead, with its longer tail now stretching out toward the higher, larger possible values of x.

Explain the hypergeometric distribution, including how it differs from the binomial distribution, state its probability formula and its mean and variance, and explain when each distribution should be used.

While the binomial distribution elegantly handles situations where repeated trials are genuinely independent of one another, an enormous number of real-world sampling situations instead involve drawing items one after another from some fixed, limited collection without ever putting each drawn item back before drawing the next one -- picking playing cards from a single deck without reshuffling them back in, selecting members for a committee from a fixed pool of candidates, or pulling components from a finite batch to test for defects, are all everyday examples of exactly this kind of situation. This particular style of sampling breaks one of the four essential conditions required for a genuinely binomial experiment, namely the strict requirement that the probability of success must remain perfectly constant across every single trial, because as soon as one item has been removed from a finite population and is not returned before the next draw, the remaining composition of that population has necessarily and permanently changed, which in turn shifts the probability of success on whatever draw comes immediately afterward. An experiment specifically built around drawing a genuinely random sample without replacement from some finite population is formally called a hypergeometric experiment, and like the binomial experiment discussed earlier, it too is fully characterized by exactly four defining properties, though these four properties differ from the binomial case in several crucially important respects. First, exactly as with the binomial experiment, the hypergeometric experiment must still be repeated some fixed, predetermined number of times, which is conventionally denoted n. Second, and here is where the two experiment types begin to diverge sharply, the successive trials of a hypergeometric experiment are inherently dependent on one another rather than independent, precisely because each individual draw permanently and irreversibly alters what specifically remains available to be drawn on every subsequent trial. Third, following directly and logically from that same dependency between trials, the probability of success is explicitly allowed to vary meaningfully from one trial to the very next trial, rather than staying artificially fixed and unchanging throughout the entire experiment as the binomial model insists it must. Fourth and finally, exactly as with the binomial case, each individual trial outcome must still be cleanly classifiable as either a success or a failure, with no other possible categories or classifications permitted. To fully specify any particular hypergeometric distribution mathematically, we generally begin by defining three separate quantities: N, representing the total overall size of the entire finite population being sampled from; k, representing how many of those N total items specifically qualify as successes within that same population; and n, representing the specific size of the particular sample being randomly drawn out from that larger population, all without any form of replacement whatsoever as items are successively drawn. Given these three carefully defined quantities, the probability of observing exactly x successes appearing within that particular random sample of size n is then formally given by the hypergeometric probability formula, P(X=x) equals the quantity C(k,x) multiplied by C(N-k, n-x), with that entire product then divided by C(N,n). The underlying logical structure embedded within this particular formula makes excellent intuitive sense once each individual piece is examined carefully in turn: the numerator's first term, C(k,x), counts every possible distinct way of choosing exactly x specific successes out from the total pool of k available successes present somewhere within the overall population; the numerator's second term, C(N-k, n-x), separately counts every possible distinct way of simultaneously choosing the remaining (n-x) failures needed to complete the sample, drawn out specifically from the (N-k) failures present within the broader population; and the single denominator term, C(N,n), simply counts the total number of ways that any sample of size n whatsoever, containing absolutely any conceivable mixture of successes and failures, could theoretically be selected from that same overall population of N total items, thereby correctly normalizing the entire calculated probability. Just as the binomial distribution offers convenient shortcut formulas for its mean and variance that avoid the tedium of constructing an entire probability table from scratch, the hypergeometric distribution similarly offers its own comparable shortcut formulas, though these particular formulas carry one additional term not present anywhere in the binomial case. The mean of the hypergeometric distribution is given simply by n.k divided by N, which bears an unmistakably close conceptual resemblance to the familiar binomial mean formula n.p, since the ratio k/N here essentially plays the exact same conceptual role that p plays over in the binomial formula, representing the underlying overall proportion of successes present within the total population being sampled from. The variance of the hypergeometric distribution, however, is given by the more elaborate expression (n.k/N) multiplied by (1 minus k/N), which itself is then further multiplied by an additional correction term, (N-n) divided by (N-1). This particular extra multiplicative term, conventionally referred to as the finite population correction factor, exists precisely because sampling without any replacement from a strictly finite, limited population mechanically produces systematically less overall variability in the resulting outcomes than would otherwise occur under comparable independent sampling with replacement, and this specific correction factor mathematically adjusts the variance downward in order to properly and accurately account for that underlying structural difference. In practical terms, deciding which of these two closely related distributions is the genuinely correct one to apply to any particular sampling problem essentially always comes down to answering one single, deceptively simple diagnostic question: is the sampling actually being conducted with replacement, or effectively from some sufficiently large, practically unlimited population where removing individual items makes no meaningful difference to the underlying probabilities involved -- in which case the binomial distribution remains the entirely appropriate choice -- or is the sampling instead being conducted explicitly without replacement from some comparatively small, genuinely finite population, where each individual item removed does meaningfully and measurably shift the probabilities that govern every subsequent draw -- in which case the hypergeometric distribution becomes the distribution that must instead be used in order to arrive at correct and reliable probability calculations.

Multiple Choice Questions (MCQs)

A trial with exactly two possible outcomes, success and failure, is called a: (A) Binomial experiment (B) Bernoulli trial (C) Hypergeometric trial (D) Random trial

Correct answer: (B) Bernoulli trial. A Bernoulli trial has exactly two outcomes with constant success probability and independence between trials.

Which of these is NOT a property of a binomial experiment? (A) Fixed number of trials (B) Constant probability of success (C) Independent trials (D) Probability of success changes each trial

Correct answer: (D) Probability of success changes each trial. A changing probability of success is a property of the hypergeometric experiment, not the binomial experiment.

The binomial probability formula is: (A) P(X=x) = np (B) P(X=x) = C(n,x) p^x q^(n-x) (C) P(X=x) = npq (D) P(X=x) = p/q

Correct answer: (B) P(X=x) = C(n,x) p^x q^(n-x). P(X=x) = C(n,x) p^x q^(n-x) gives the probability of exactly x successes in n independent trials.

A binomial distribution is symmetric when: (A) p = 0 (B) p = 1 (C) p = q = 0.5 (D) n = 1

Correct answer: (C) p = q = 0.5. When p equals q (both 0.5), the binomial distribution is perfectly symmetric.

If p > 0.5 in a binomial distribution, the distribution is: (A) Symmetric (B) Positively skewed (C) Negatively skewed (D) Uniform

Correct answer: (C) Negatively skewed. When p exceeds 0.5, success is more likely, and the distribution becomes negatively skewed.

The mean of a binomial distribution is given by: (A) npq (B) np (C) sqrt(npq) (D) n/p

Correct answer: (B) np. The mean of the binomial distribution is E(X) = np.

The variance of a binomial distribution is given by: (A) np (B) npq (C) nq (D) sqrt(np)

Correct answer: (B) npq. The variance of the binomial distribution is Var(X) = npq.

A hypergeometric experiment involves sampling: (A) With replacement from an infinite population (B) Without replacement from a finite population (C) With replacement from a finite population (D) Without any population

Correct answer: (B) Without replacement from a finite population. The hypergeometric distribution specifically applies to sampling without replacement from a finite population.

The hypergeometric probability formula is: (A) P(X=x) = C(n,x)p^x q^(n-x) (B) P(X=x) = [C(k,x).C(N-k,n-x)] / C(N,n) (C) P(X=x) = nk/N (D) P(X=x) = np

Correct answer: (B) P(X=x) = [C(k,x).C(N-k,n-x)] / C(N,n). The hypergeometric formula divides the product of two combinations by C(N,n).

The finite population correction factor in the hypergeometric variance formula is: (A) np (B) npq (C) (N-n)/(N-1) (D) k/N

Correct answer: (C) (N-n)/(N-1). The factor (N-n)/(N-1) adjusts the variance to account for sampling without replacement from a finite population.

Quick Revision Summary

- Binomial experiment: fixed n trials, constant p, independent trials, success/failure outcomes
- Binomial formula: P(X=x) = C(n,x) p^x q^(n-x)
- Binomial shape: p=0.5 symmetric | p>0.5 negatively skewed | p<0.5 positively skewed
- Binomial mean = np; Binomial variance = npq; SD = sqrt(npq)

- Binomial frequency distribution: f(X=x) = N.P(X=x) for N repetitions of the experiment
- Hypergeometric experiment: fixed n, dependent trials, changing probability of success, finite population
- Hypergeometric formula: P(X=x) = [C(k,x).C(N-k,n-x)] / C(N,n)
- Hypergeometric mean = nk/N; variance = (nk/N)(1-k/N)((N-n)/(N-1)) — includes finite population correction
- Use binomial when sampling WITH replacement (or huge population); use hypergeometric WITHOUT replacement from a small finite population

Exam Tips

- Always check the four binomial conditions before applying the binomial formula — 'with replacement' or 'independent' are the key trigger words
- 'Without replacement' from a small finite population is the trigger phrase for hypergeometric — don't default to binomial
- For 'at least' or 'at most' binomial problems, list out and sum the individual P(X=x) terms rather than trying to shortcut it
- Memorize mean=np and variance=npq for binomial — they save enormous time versus building a full probability table
- In hypergeometric problems, carefully identify N (population), k (successes in population), and n (sample size) before plugging into the formula