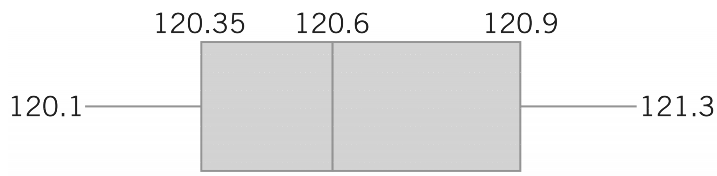
Answers to Exercise Problems

Q4.1 [Ans: (a) $950; (b)$ 302.77]

Q4.2 [Ans: (a) $10.2; (b)$ 35.06]

Q4.3 [Ans: (1) The diameters appear to be centered around 55 to 56 mm, which indicates good consistency in the manufacturing process. There is a fairly symmetric distribution around these values. (2) There are no extreme values (outliers) that deviate significantly from the rest of the data, suggesting that the process is stable and in control.]

Q4.4 [Ans: Median ≈ 120.6 mm; $Q_1 \approx 120.35$ mm; $Q_3 \approx 120.9$ mm; $IQR \approx 0.55$ mm]



Q4.5 [Ans: ≈ 0.9742]

Q4.6 [Ans: ≈ 0.3333]

Q4.7 [Ans: ≈ 359 ml]

Q4.8 [Ans: $\approx 95.5\%, 104.5\%$]

5 Discrete Distributions

Discrete Distributions

Distribution	Use Case in Quality Engineering
Geometric	Determining the probability of the *first occurrence* of a defect on the nth trial; estimating the number of trials required to encounter the *first defect* in a production line.
Hypergeometric	Quality control sampling *without replacement* (e.g., finding the probability of a certain number of defects in a sample from a batch); estimating defect rates in small lot sizes where items are *not replaced after testing.*
Binomial	Estimating the probability of a given number of defective items in a batch; analyzing the *success/failure outcomes* of a series of independent tests (e.g., pass/fail in quality inspections).
Negative Binomial	Modeling the *number of trials* needed to achieve a *specified number of successes;* analyzing the failure rates and determining the process stability over repeated tests.
Poisson	Modeling the number of defects in a process over a *fixed period or space;* assessing the probability of a certain number of events (e.g., defects) occurring within a *given time-frame.*

Hypergeometric Distribution

Suppose that there is a finite population consisting of N items. D ($D \leq N$) of these items fall into a class of interest. A random sample of n items is selected from the population N without replacement. It is observed that X number of items in the sample fall into the class of interest D. Then X is a hypergeometric random variable.

- The hypergeometric probability mass function (*PMF*) can be written as

$$P(X = x) = \frac{{}^{D}C_x \; {}^{N-D}C_{n-x}}{{}^{N}C_n}, \quad x = 0, 1, 2, \cdots, \min(n, D)$$

- The *mean* and *variance* of the distribution are

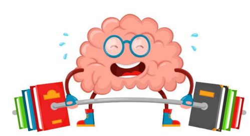
$$\mu = \frac{nD}{N}$$
$$\sigma^2 = \frac{nD}{N}\left(1 - \frac{D}{n}\right)\left(\frac{N-n}{N-1}\right)$$

Hypergeometric Distribution (cont'd)

In the context of quality control, the hypergeometric distribution is the appropriate probability model for selecting a random sample of n items without replacement from a lot of N items of which D are nonconforming or defective.

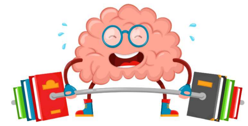
- A *random sample* is a sample that has been selected in such a way that all possible samples have an *equal chance* of being chosen.
- In these applications, X usually represents the number of nonconforming items found in the sample.

Exercise Problems



Q5.1 In quality engineering, particularly in quality control and assurance, it's common to inspect a sample of items from a production lot to determine if the lot meets certain quality standards. In this scenario, you have a lot of 100 items, which represents the entire batch produced. Among these 100 items, 5 items are nonconforming (i.e., they do not meet the required quality standards). A sample of 10 items is selected at random from this lot without replacement. Determine the probability that this sample contains one or fewer nonconforming items.

Exercise Problems



Q5.1 In quality engineering, particularly in quality control and assurance, it's common to inspect a sample of items from a production lot to determine if the lot meets certain quality standards. In this scenario, you have a lot of 100 items, which represents the entire batch produced. Among these 100 items, 5 items are nonconforming (i.e., they do not meet the required quality standards). A sample of 10 items is selected at random from this lot without replacement. Determine the probability that this sample contains one or fewer nonconforming items.

Q5.2 Suppose you have a manufacturing process that produces electronic components. Each component produced has a 0.5% chance of being defective. To ensure quality, a lot of 1000 components is produced, and a quality engineer decides to randomly sample 50 components from this lot without replacement. Given this setup, calculate the probability that the sample contains exactly 2 defective components.

Binomial Distribution

Consider a process that consists of a sequence of n independent trials, i.e., the outcome of each trial does not depend in any way on the outcome of previous trials. When the outcome of each trial is either a *success* or a *failure,* the trials are called *Bernoulli trials.*

- If the probability of *success* on any trial is p, then the number of *successes* X in n Bernoulli trials has the binomial distribution with parameters $n \geq 0$ and $0 < p < 1$.

- The binomial *PMF* can be written as

$$P(X = x) = {}^{n}C_{x} p^{x} (1-p)^{n-x}, \quad x = 0, 1, 2, \cdots, n$$

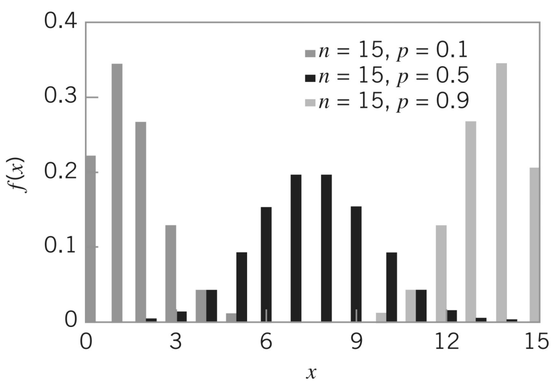
- The *mean* and *variance* of the distribution are

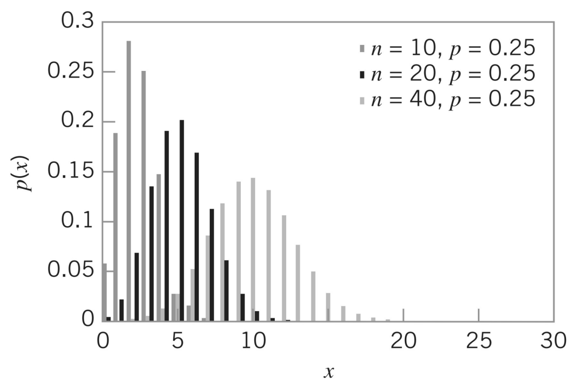
$$\mu = np, \quad \sigma^2 = np(1-p)$$

Binomial Distribution (cont'd)

The binomial distribution is used frequently in quality engineering. It is the appropriate probability model for sampling from an infinitely large population, where p represents the *fraction of defective or nonconforming items* in the population.

- In these applications, X usually represents the number of nonconforming items found in a random sample of n items.

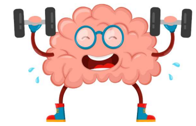




Exercise Problems

Q5.3 A manufacturing company produces electronic components, where each component has a 3% chance of being defective. The quality control team randomly selects 15 components from a batch for inspection. What is the probability of obtaining two or fewer nonconforming items in the sample?

Exercise Problems



Q5.3 A manufacturing company produces electronic components, where each component has a 3% chance of being defective. The quality control team randomly selects 15 components from a batch for inspection. What is the probability of obtaining two or fewer nonconforming items in the sample?

Q5.4 A company produces a batch of 1000 electronic devices, and historical data shows that 988 of the devices pass the quality test. The company decides to test a sample of 20 devices from this batch. (a) What is the probability that all 20 devices pass the quality test? (b) What is the probability that at least 18 devices pass the quality test?

Poisson Distribution

Consider X be the number of events occurring in a fixed interval of time or space, given a constant average rate of occurrence $\lambda > 0$. Here X represents a Poisson random variable.

- The Poisson *PMF* can be written as

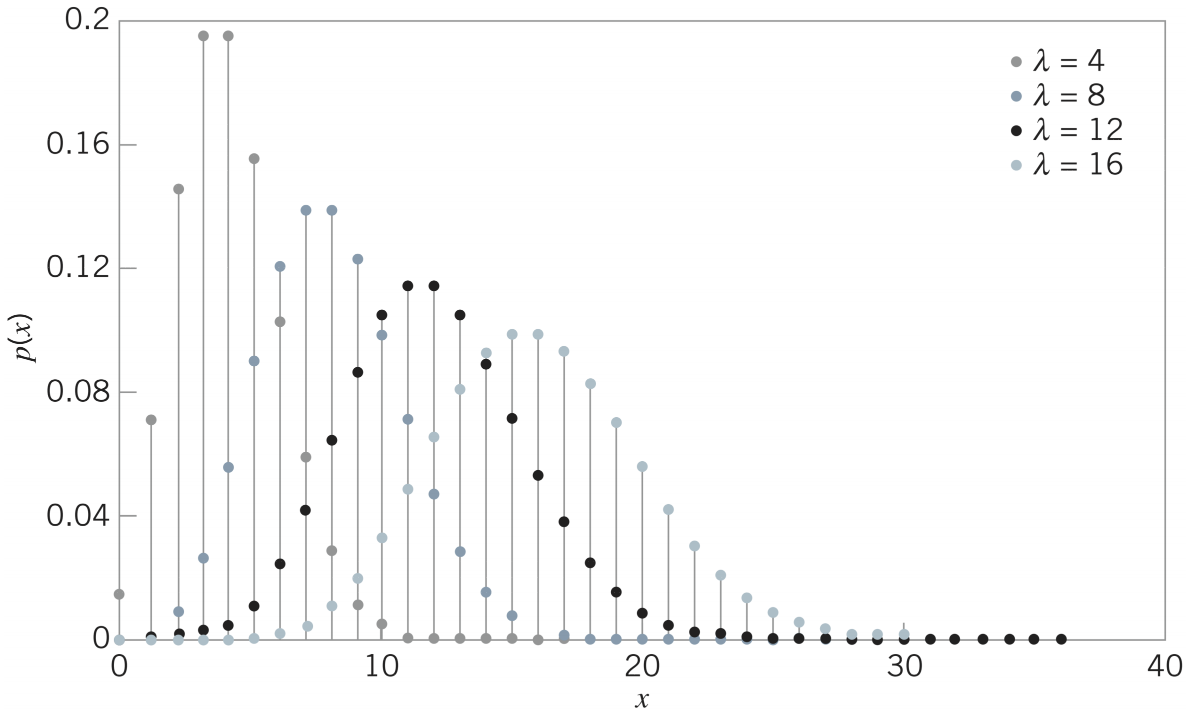
$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \cdots$$

- The *mean* and *variance* of the distribution are

$$\mu = \lambda, \quad \sigma^2 = \lambda$$

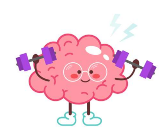
- Any random phenomenon that occurs on a per unit (or per unit volume, per unit time, etc.) basis is often well approximated by the Poisson distribution.

Poisson Distribution (cont'd)



© 2025 Indranil Hazra

Exercise Problems



Q5.5 A semiconductor device manufacturing process has a known average of 4 wire-bonding defects per unit. The number of wire-bonding defects per unit is assumed to follow a Poisson distribution. What is the probability that a randomly selected semiconductor device will have two or fewer wire-bonding defects?

Exercise Problems

Q5.5 A semiconductor device manufacturing process has a known average of 4 wire-bonding defects per unit. The number of wire-bonding defects per unit is assumed to follow a Poisson distribution. What is the probability that a randomly selected semiconductor device will have two or fewer wire-bonding defects?

Q5.6 A software quality assurance (QA) team is testing a large software system. The number of bugs found in any given module of the software follows a Poisson distribution with a mean of 5 bugs per module. The system contains 10 independent modules, and the QA team will only release the software if no more than 3 modules contain 7 or more bugs. (a) What is the probability that a randomly selected module contains 7 or more bugs? (b) What is the probability that no more than 3 out of the 10 modules contain 7 or more bugs? (c) Based on this probability, should the QA team release the software?

Negative Binomial Distribution

The negative binomial distribution, like the binomial distribution, has its basis in Bernoulli trials. Consider a sequence of independent trials, each with probability of success p, and let X denote the trial on which the rth success occurs. Then, X represents a negative binomial random variable.

– The negative binomial *PMF* can be written as

$$P(X = x) = {}^{x-1}C_{r-1} p^r (1-p)^{x-r}, \quad x = r, r+1, r+2, \cdots$$

– The *mean* and *variance* of the distribution are

$$\mu = \frac{r}{p}, \quad \sigma^2 = \frac{r(1-p)}{p^2}$$

Exercise Problems

Q5.7 A quality control inspector is checking items produced on a manufacturing line. The probability that an item is defective is 0.05. The inspector will keep checking items until they find the 3rd defective one. (a) What is the probability that the inspector will need to inspect exactly 10 items to find the 3rd defective one? (b) What is the probability that the inspector will need to inspect at most 5 items to find the 3rd defective one?

[5] Discrete Distributions • § Negative Binomial & Geometric Distributions

© 2025 Indranil Hazra

Exercise Problems

Q5.7 A quality control inspector is checking items produced on a manufacturing line. The probability that an item is defective is 0.05. The inspector will keep checking items until they find the 3rd defective one. (a) What is the probability that the inspector will need to inspect exactly 10 items to find the 3rd defective one? (b) What is the probability that the inspector will need to inspect at most 5 items to find the 3rd defective one?

Q5.8 A factory operates several identical machines, each with a 2% chance of failure on any given day. The factory management wants to schedule preventive maintenance after a certain number of failures. Specifically, they will perform maintenance after the 5th machine failure. (a) What is the probability that the 5th machine failure will occur on the 30th day? (b) What is the expected number of days until the 5th machine failure occurs? (c) What is the standard deviation of number of days until the the 5th machine failure occurs?

Dr. Indranil Hazra RE30003 @ IIT Kharagpur Spring 2025 92

Geometric Distribution

A useful *special case* of the negative binomial distribution is if $r = 1$, in which case we have the geometric distribution. It is the distribution of the number of *Bernoulli trials until the first success.*

- The geometric *PMF* can be written as

$$P(X = x) = p(1-p)^{x-1}, \quad x = 1, 2, \cdots$$

- The *mean* and *variance* of the distribution are

$$\mu = \frac{1}{p}, \quad \sigma^2 = \frac{(1-p)}{p^2}$$

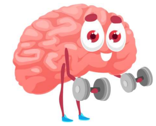
Geometric Distribution (cont'd)

- Because the sequence of Bernoulli trials are independent, the count of the number of trials until the next success can be started from anywhere without changing the probability distribution.

- The geometric distribution shows the *lack of memory.* This property implies that the system being modeled does not show any form of wear or degradation.

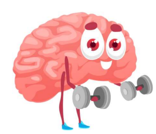
- The *negative binomial random variable can be defined as the sum of geometric random variables.* That is, the sum of r geometric random variables each with parameter p is a negative binomial random variable with parameters p and r.

Exercise Problems



Q5.9 A software testing team is evaluating a new system's performance and finds that, on average, a particular type of memory latency error occurs with a probability of 0.1 during testing. The team is interested in the number of trials (test runs) needed to encounter the first memory latency error. (a) What is the probability that the team will encounter the first memory latency error on the 5th test run? (b) What is the probability that the team will encounter the first memory latency error within the first 3 test runs?

Exercise Problems



Q5.9 A software testing team is evaluating a new system's performance and finds that, on average, a particular type of memory latency error occurs with a probability of 0.1 during testing. The team is interested in the number of trials (test runs) needed to encounter the first memory latency error. (a) What is the probability that the team will encounter the first memory latency error on the 5th test run? (b) What is the probability that the team will encounter the first memory latency error within the first 3 test runs?

Q5.10 In a quality control process, a defective component occurs with a probability of 0.2 during each inspection. An engineer wants to understand the number of inspections required to find a certain number of defective components. Assume that the number of inspections until a defective component is found follows a geometric distribution with $p = 0.2$, where p is the probability of finding a defective component in a given inspection. (a) What is the expected number of inspections required to find the 4th defective component? (b) If the engineer needs to find 3 defective components, what is the variance in the total number of inspections required?

Answers to Exercise Problems

Q5.1 [Ans: $p \approx 0.9231$]

Q5.2 [Ans: $p \approx 0.0212$]

Q5.3 [Ans $p \approx 0.9906$]

Q5.4 [Ans: $(a)\ p \approx 0.7855,\ \ (b)\ p \approx 0.9983$]

Q5.5 [Ans $p \approx 0.2381$]

Q5.6 [Ans: $(a)\ p \approx 0.2378, (b)\ p \approx 0.8036$]

Q5.7 [Ans $(a)\ p \approx 0.0031,\ \ (b)\ p \approx 0.0012$]

Q5.8 [Ans: $(a)\ p \approx 4.6 \times 10^{-5},\ \ (b)\ 250$ days, $(b)\ 111$ days]

Q5.9 [Ans $(a)\ p \approx 0.0656,\ \ (b)\ p \approx 0.271$]

Q5.10 [Ans: $(a)\ 20,\ \ (b)\ 60$]