Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

EEU33E03: Probability and Statistics

Lecture 6: Discrete Random Variables and Probability Distributions

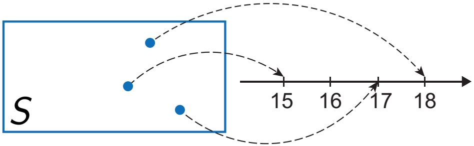
Arman Farhang (arman.farhang@tcd.ie)

2023

Outline

Random Variables

A **random variable** is a function that assigns a real number to each outcome in the sample space of a random experiment.



A random variable is represented by an uppercase letter such as X and the particular value of the corresponding random variable is denoted by a lowercase letter.

The notation $X(s) = x$ means that x is the value associated with the outcome s by the random variable X.

More formally, a random variable is a function whose domain is the sample space and whose range is the set of real numbers.

Types of Random Variables

Continuous Random Variables: A continuous random variable is a random variable with an interval (either finite or infinite) of real numbers for its range while no possible value of the random variable has positive probability, i.e., $P(X = x_0) = 0$.
Some examples of the continuous random variables include:

- electrical current, length, pressure, temperature, time, voltage, weight.

Discrete Random Variables: A discrete random variable is a random variable with a finite (or countably infinite) range.
Some examples of the discrete random variables include:

- number of transmitted bits received in error in a communication system, number of scratches on a surface, proportion of defective parts among 1000 tested.

Types of Random Variables

Example 1: Computer chips often contain surface imperfections. For a certain type of computer chip, 9% contain no imperfections, 22% contain 1 imperfection, 26% contain 2 imperfections, 20% contain 3 imperfections, 12% contain 4 imperfections, and the remaining 11% contain 5 imperfections. Let X represent the number of imperfections in a randomly chosen chip. What are the possible values for X? Is X discrete or continuous? Find $P(X = x)$ for each possible value x.

Types of Random Variables

Example 1: Computer chips often contain surface imperfections. For a certain type of computer chip, 9% contain no imperfections, 22% contain 1 imperfection, 26% contain 2 imperfections, 20% contain 3 imperfections, 12% contain 4 imperfections, and the remaining 11% contain 5 imperfections. Let X represent the number of imperfections in a randomly chosen chip. What are the possible values for X? Is X discrete or continuous? Find $P(X = x)$ for each possible value x.

Solution: The possible values for X are the integers $0, 1, 2, 3, 4, 5$. The random variable X is discrete, as it takes on only integer values. Nine percent of the outcomes in the sample space are assigned the value 0. Thus, $P(X = 0) = 0.09$, $P(X = 1) = 0.22$, $P(X = 2) = 0.26$, $P(X = 3) = 0.20$, $P(X = 4) = 0.12$, and $P(X = 5) = 0.11$.

Discrete Random Variables

Probability distribution and probability mass function

The list of possible values for the random variable X in the previous example, along with the probabilities for each, provide a complete description of the population from which X is drawn.

This description is called the **probability mass function** or the **probability distribution**.

Definition

For a discrete random variable X with possible values $x_1, x_2, \ldots, x_n$, a probability mass function (PMF) is a function such that

1. $f(x_i) = P(X = x_i)$,
2. $f(x_i) \geq 0$,
3. $\sum_{i=1}^{n} f(x_i) = 1$.

Cumulative Distribution Function

In Example 1, we might be interested in the probability of three or fewer imperfections, i.e., $P(X \leq 3)$.

Since the event $X \leq 3$ is the union of the events $\{X = 0\}$, $\{X = 1\}$, $\{X = 2\}$, and $\{X = 3\}$,

$$\begin{aligned} P(X \leq 3) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= 0.09 + 0.22 + 0.26 + 0.20 = 0.77. \end{aligned}$$

If we have a function that specifies the probability that a random variable is less than or equal to a given value, we can use this function to calculate the probability of different events, e.g., $P(X = 3) = P(X \leq 3) - P(X \leq 2)$.

This function is called the **cumulative distribution function**.

Cumulative Distribution Function

Definition

The **cumulative distribution function (CDF)** of a discrete random variable X, denoted as $F(x)$, is

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i). \tag{1}$$

This function satisfies the following properties.

1. $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i) = \sum_{x_i \leq x} P(X = x_i)$,
2. $0 \leq F(x) \leq 1$,
3. If $x \leq y$, then $F(x) \leq F(y)$.

Example

Example 2: The system below contains two components, A and B, connected in series. The system will function only if both components function. The probability that A and B function is given by $P(A) = 0.96$, and $P(B) = 0.92$, respectively. Assume that A and B function independently. Define a random variable X for the system and find the corresponding PMF and CDF functions.

Example

Example 2: The system below contains two components, A and B, connected in series. The system will function only if both components function. The probability that A and B function is given by $P(A) = 0.96$, and $P(B) = 0.92$, respectively. Assume that A and B function independently. Define a random variable X for the system and find the corresponding PMF and CDF functions.

Solution: Let us define E_A as the event that the first device functions and E_B as the event that the second device functions. There are 3 possible outcomes, (i) both devices function, (ii) only one of the devices functions, and (iii) none of the devices function, which we map to $X = 0$, $X = 1$, and $X = 2$, respectively. Thus, the PMF is obtained by finding the probabilities $P(X = 0)$, $P(X = 1)$, and $P(X = 2)$. Using these probabilities, we can also calculate the CDF.

Example

Solution (continued):

$P(X=0) = P(E_A \cap E_B) = (0.96)(0.92) = 0.8832$

$P(X=1) = P(E_A^C \cap E_B) + P(E_A \cap E_B^C) = (1-0.96)(0.92) + (0.96)(1-0.92) = 0.1136$

$P(X=2) = P(E_A^C \cap E_B^C) = (1-0.96)(1-0.92) = 0.0032$

$$f(x) = \begin{cases} 0.8832, & X=0 \\ 0.1136, & X=1 \\ 0.0032, & X=2 \end{cases}$$

Cumulative distribution function $F(x)$ can be obtained as

$F(X=0) = P(X \leq 0) = 0.8832$

$F(X=1) = P(X \leq 1) = 0.8832 + 0.1136 = 0.9968$

$F(X=2) = P(X \leq 2) = 0.8832 + 0.1136 + 0.0032 = 1$

$$F(x) = \begin{cases} 0.8832, & X=0 \\ 0.9968, & X=1 \\ 1, & X=2 \end{cases}$$

Mean and Variance

Definition

The **mean**, **expected value** or **expectation** of a discrete random variable X, denoted as μ or $E(X)$, is given by

$$\mu = E(X) = \sum_x xf(x). \tag{2}$$

The **variance** of X, denoted as σ^2 or $V(X)$, is given by

$$\sigma^2 = V(X) = E\{(X-\mu)^2\} = \sum_x (x-\mu)^2 f(x) = \sum_x x^2 f(x) - \mu^2. \tag{3}$$

The **standard deviation** is the square root of the variance, i.e., $\sigma = \sqrt{\sigma^2}$.

The Expected Value of a Function

Sometimes, we are interested in finding the expected value of a function of the random variable X, i.e., $h(X)$.

Example 3: The cost of a certain vehicle diagnostic test depends on the number of cylinders X in the vehicle's engine. Suppose the cost function is given by $h(X) = \frac{1}{2}X^2 + 3X + 20$. Since X is a random variable, so is $Y = h(X)$. The pmf of X and derived pmf of Y are as follows:

x	4	6	8
f(x)	0.5	0.3	0.2

$\longrightarrow$

y	40	56	76
f(y)	0.5	0.3	0.2

$$E(Y) = E[h(X)] = \sum_y yf(y) = (40)(0.5) + (56)(0.3) + (76)(0.2)$$

$$= h(4) \times (0.5) + h(6) \times (0.3) + h(8) \times (0.2) = \sum_x h(x)f(x)$$

Expected Value Properties

Properties of Expected Value

1. The function of interest is often a linear function, $h(X) = aX + b$. In this case, $E[h(X)]$ can be easily obtained as

$$E(aX + b) = aE(X) + b \qquad \text{OR} \qquad \mu_{aX+b} = a\mu + b.$$

2. Expected value of a given constant a is itself, i.e., $E(a) = a$.
3. Expectation of a positive random variable is always positive.

Proofs are left as exercise.

Properties of Variance

Properties of Variance

1. For a linear function of interest, $h(X) = aX + b.$, $V[h(X)]$ can be easily obtained.

$$V(aX + b) = \sigma^2_{aX+b} = a^2 V(x) = a^2 \sigma^2_X \qquad \text{and} \qquad \sigma_{aX+b} = |a|\sigma_X.$$

2. Variance of a given constant a is zero, i.e., $V(a) = a$.

Proofs are left as exercise.

Distributions of Discrete Random Variables

The Uniform Distribution

A random variable X has a **discrete uniform distribution** if each of the n values in its range, say, $x_1, x_2, \ldots, x_n$, has equal probability.

$$f(x_i) = \frac{1}{n} \tag{4}$$

Suppose X is a discrete uniform random variable on the consecutive integers $a, a+1, \ldots, b$, for $a \leq b$. The mean of X is

$$\mu = E(X) = \frac{b+a}{2}, \tag{5}$$

and the variance of X is

$$\sigma^2 = \frac{(b-a+1)^2 - 1}{12}. \tag{6}$$

Distributions of Discrete Random Variables

The Bernoulli Distribution

The Bernoulli Distribution ($X \sim \mathrm{Bernoulli}(p)$)

X has a $0-1$ distribution with parameter p, we write $X \sim \mathrm{Bernoulli}(p)$, if

$$P(X=0) = 1-p,$$
$$P(X=1) = p.$$

If $X \sim \mathrm{Bernoulli}(p)$, then its mean and variance are

$$\mu_X = p, \tag{7}$$

and

$$\sigma_X^2 = p(1-p), \tag{8}$$

Proofs are left as exercise.

Distributions of Discrete Random Variables

The Binomial Distribution

Considering total of n Bernoulli trials such that

1. The trials are independent,
2. Each trial has the same success probability p, and
3. X is the number of trials that result in a success,

then the random variable X has a **binomial distribution** with parameters n and p, denoted as $X \sim \mathrm{Bin}(n, p)$. The probability mass function for this random variable is

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}. \tag{9}$$

The name of the distribution is obtained from the *binomial expansion* for constants a and b, i.e., $(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}$.

The Binomial Distribution: Example

Example 4: A fair die is rolled 5 times. Find (i) the probability that 2 sixes come up: $P(A)$; (ii) the probability that at least one six comes up: $P(B)$, (iii) the probability of getting least two sixes, given that at least one six comes up: $P(C|B)$.

The Binomial Distribution: Example

Example 4: A fair die is rolled 5 times. Find (i) the probability that 2 sixes come up: $P(A)$; (ii) the probability that at least one six comes up: $P(B)$, (iii) the probability of getting least two sixes, given that at least one six comes up: $P(C|B)$.

Solution: The probability of getting a 6 in one roll of the die is $\frac{1}{6}$. The random variable X that represents the number of sixes when we roll the die 5 times has a $\mathrm{Bin}(5, \frac{1}{6})$ distribution.

(i) $$P(A) = P(X = 2) = \binom{5}{2}\left(\frac{1}{6}\right)^2\left(1 - \frac{1}{6}\right)^{(5-2)} = \frac{5!}{2!3!}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^3 \approx 0.1607$$

The Binomial Distribution: Example

Example 4: A fair die is rolled 5 times. Find (i) the probability that 2 sixes come up: $P(A)$; (ii) the probability that at least one six comes up: $P(B)$, (iii) the probability of getting least two sixes, given that at least one six comes up: $P(C|B)$.

Solution: The probability of getting a 6 in one roll of the die is $\frac{1}{6}$. The random variable X that represents the number of sixes when we roll the die 5 times has a $\mathrm{Bin}(5, \frac{1}{6})$ distribution.

(i)
$$P(A) = P(X = 2) = \binom{5}{2}\left(\frac{1}{6}\right)^2\left(1 - \frac{1}{6}\right)^{(5-2)} = \frac{5!}{2!3!}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^3 \approx 0.1607$$

(ii)
$$P(B) = P(X \geq 1) = \sum_{i=1}^{5}\binom{5}{i}\left(\frac{1}{6}\right)^i\left(\frac{5}{6}\right)^{(5-i)}$$
$$= 1 - P(X < 1) = 1 - P(X = 0) = 1 - \binom{5}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^{(5-0)} \approx 0.5981$$

The Binomial Distribution: Example

Solution (Continued):

(iii)

$$P(C|B) = P(X \geq 2|X \geq 1) = \frac{P(X \geq 2 \cap X \geq 1)}{P(X \geq 1)} = \frac{P(X \geq 2)}{P(X \geq 1)}$$

Since $P(X \geq 1) = P(X = 1) + P(X \geq 2)$, thus, $P(X \geq 2) = P(X \geq 1) - P(X = 1)$.
Hence,

$$P(X \geq 2) = 0.5981 - \binom{5}{1}\left(\frac{1}{6}\right)^{1}\left(\frac{5}{6}\right)^{(5-1)} \approx 0.1962.$$

Consequently,

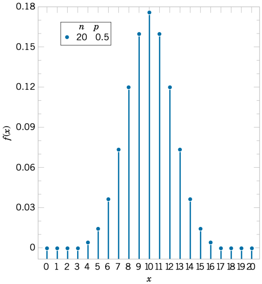
$$P(C|B) = P(X \geq 2|X \geq 1) = \frac{0.1962}{0.5981} \approx 0.3281.$$

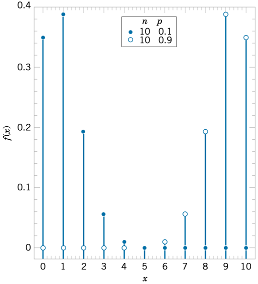
Distributions of Discrete Random Variables
The Binomial Distribution

Examples of binomial distributions are shown in the above figure on the right.

For a fixed n, the distribution becomes more symmetric as p increases from 0 to 0.5 or decreases from 1 to 0.5.

For a fixed p, the distribution becomes more symmetric as n increases.





If $X \sim \mathrm{Bin}(n, p)$, then its mean and variance are

$$\mu_X = np \qquad \sigma_X^2 = np(1-p)$$

Distributions of Discrete Random Variables

The Geometric Distribution

Considering a sequence of Bernoulli trials, such that

1. The trials are independent,
2. Each trial has the same success probability p, and
3. The random variable X represents the number of trials up to and including the first success,

then the random variable X has a **geometric distribution** with parameter p, and it is denoted as $X \sim \mathrm{Geom}(p)$. The probability mass function for the geometric random variable X is

$$f(x) = (1-p)^{x-1}p. \tag{10}$$

If $X \sim \mathrm{Geom}(p)$, then its mean and variance are

$$\mu_X = 1/p \qquad \sigma_X^2 = (1-p)/p^2$$

Distributions of Discrete Random Variables

The Negative Binomial Distribution

Considering a sequence of Bernoulli trials, such that

1. The trials are independent,
2. Each trial has the same success probability p, and
3. The random variable X represents the number of trials up to and including the r^{th} success,

then the random variable X has a **negative binomial distribution** with parameters r and p, and it is denoted as $X \sim \mathrm{NB}(r, p)$ and the probability mass function

$$f(x) = \binom{x-1}{r-1}(1-p)^{x-r}p^r. \tag{11}$$

If $X \sim \mathrm{NB}(r, p)$, then its mean and variance are

$$\mu_X = r/p \qquad \sigma_X^2 = r(1-p)/p^2$$

Distributions of Discrete Random Variables

The Hypergeometric Distribution

Let us consider

1. A population or set to be sampled, with N individuals, objects, or elements, i.e., *a finite population*.

Distributions of Discrete Random Variables
The Hypergeometric Distribution

Let us consider

1. A population or set to be sampled, with N individuals, objects, or elements, i.e., *a finite population*.
2. Each individual can be characterized as a ‘success’ or a ‘failure’, and there are K number of successes in the population.
   Note: Each item that is sampled constitutes a Bernoulli trial.

Distributions of Discrete Random Variables

The Hypergeometric Distribution

Let us consider

1. A population or set to be sampled, with N individuals, objects, or elements, i.e., *a finite population*.
2. Each individual can be characterized as a 'success' or a 'failure', and there are K number of successes in the population.
   Note: Each item that is sampled constitutes a Bernoulli trial.
3. A sample of n individuals is randomly selected without replacement so that each subset of size n is equally likely to be chosen.

Distributions of Discrete Random Variables

The Hypergeometric Distribution

Let us consider

1. A population or set to be sampled, with N individuals, objects, or elements, i.e., *a finite population*.
2. Each individual can be characterized as a 'success' or a 'failure', and there are K number of successes in the population.
   Note: Each item that is sampled constitutes a Bernoulli trial.
3. A sample of n individuals is randomly selected without replacement so that each subset of size n is equally likely to be chosen.

The random variable X, denoting the number of successes in the sample, does not follow a binomial distribution, instead, it follows a distribution called **hypergeometric distribution**.

The Hypergeometric Distribution: Example

Example 5: Among 50 electronic chipsets that are manufactured, 10 are defective, and that 6 items are sampled from this set at random. Let X be the number of defective items in the sample and compute $P(X = 4)$.

The Hypergeometric Distribution: Example

Example 5: Among 50 electronic chipsets that are manufactured, 10 are defective, and that 6 items are sampled from this set at random. Let X be the number of defective items in the sample and compute $P(X = 4)$.

Solution: We need to count the total number of different groups of 6 chipsets that can be sampled from the whole population of 50.

The Hypergeometric Distribution: Example

Example 5: Among 50 electronic chipsets that are manufactured, 10 are defective, and that 6 items are sampled from this set at random. Let X be the number of defective items in the sample and compute $P(X = 4)$.

Solution: We need to count the total number of different groups of 6 chipsets that can be sampled from the whole population of 50.

$$\binom{50}{6} = \frac{50!}{6!(50-6)!} = 15,890,700$$

The Hypergeometric Distribution: Example

Example 5: Among 50 electronic chipsets that are manufactured, 10 are defective, and that 6 items are sampled from this set at random. Let X be the number of defective items in the sample and compute $P(X = 4)$.

Solution: We need to count the total number of different groups of 6 chipsets that can be sampled from the whole population of 50.

$$\binom{50}{6} = \frac{50!}{6!(50-6)!} = 15,890,700$$

Then, we find the number of the combinations of 6 chipsets that contain exactly 4 defectives. To this end, first, we select 4 items from the 10 defective chipsets, and then, we select 2 items from the 40 flawless ones. Combinations of 6 chipsets that can be made up of 4 defective and 2 non-defective ones can be obtained as $\binom{10}{4}\binom{40}{2} = 210 \times 780 = 163800$. Therefore, $P(X = 4) = \frac{163800}{15890700} = 0.0103$.

Distributions of Discrete Random Variables
The Hypergeometric Distribution

A set of N objects contains

- K objects classified as 'successes'
- $N - K$ objects classified as 'failures'

A sample of size n objects is selected randomly (without replacement) from the N objects, where $K \leq N$ and $n \leq N$. Let the random variable X denote the number of successes in the sample. Then X is a **hypergeometric random variable** and

$$f(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}, \qquad \max\{0, n+K-N\} \leq x \leq \min\{K, n\}$$

For the random variable X, considering $p = K/N$, $\mu_X = np$ and $\sigma_X^2 = np(1-p)\left(\frac{N-n}{N-1}\right)$.

Distributions of Discrete Random Variables
The Multinomial Distribution

Suppose a random experiment consists of a series of n independent trials, and

1. The result of each trial is classified into one of k classes.
2. The probability of a trial generating a result in class 1, class 2, ... , class k is constant over the trials and equal to $p_1, \ldots, p_k$, respectively.

The random variables $X_1, \ldots, X_k$ denoting the number of trials that result in class 1, class 2, ... , class k, respectively, have a multinomial distribution and the joint probability mass function is

$$P(X_1 = x_1, \ldots, X_k = x_k) = \frac{n!}{x_1!x_2!\ldots x_k} p_1^{x_k} \ldots p_k^{x_k}, \tag{12}$$

where $x_1 + \ldots + x_k = n$ snd $p_1^{x_k} + \ldots + p_k^{x_k} = 1$.

Distributions of Discrete Random Variables
The Poisson Distribution

Poison processes frequently appear in scientific studies, in connection with the occurrence of events of some kind over time or space.

Poisson distribution is as an approximation to the binomial distribution when n is large and p is small.

Example 6: Consider the transmission of n bits over a digital communication channel. Let the random variable X represent the number of erroneous bits. When the probability of error for a bit is constant, and the transmissions are independent, X has a binomial distribution. Let p denote the probability of error for a bit. Therefore,

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} = \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x},$$

where $\lambda = \mu_X = np$.

Example 6 (Continued): Now, suppose that the number of bits transmitted increases and the probability of an error decreases exactly enough that pn remains constant, i.e., $\lambda = np$ is constant. Thus,

$$\begin{aligned}
\lim_{n\to\infty} P(X=x) &= \lim_{n\to\infty} \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x} \\
&= \lim_{n\to\infty} \frac{n!}{x!(n-x)!} \lambda^x \left(\frac{1}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{-x} \left(1-\frac{\lambda}{n}\right)^n \\
&= \frac{\lambda^x}{x!} \lim_{n\to\infty} \frac{n!}{(n-x)!(n-\lambda)^x} \left(1-\frac{\lambda}{n}\right)^n \\
&= \frac{\lambda^x}{x!} \lim_{n\to\infty} \frac{(n-x+1)\cdots(n)}{(n-\lambda)^x} \left(1-\frac{\lambda}{n}\right)^n = \frac{e^{-\lambda}\lambda^x}{x!}
\end{aligned}$$

Example 7: A mass contains 10,000 atoms of a radioactive substance. The probability that a given atom will decay in a one-minute time period is 0.0002. Let X represent the number of atoms that decay in one minute.

Distributions of Discrete Random Variables
The Poisson Distribution

Example 7: A mass contains 10,000 atoms of a radioactive substance. The probability that a given atom will decay in a one-minute time period is 0.0002. Let X represent the number of atoms that decay in one minute.

Each atom can be considered as a Bernoulli trial, where decay of an atom represents 'success'. Since, X denotes the number of successes in 10,000 independent Bernoulli trials, $X \sim \mathrm{Bin}(10000, 0.0002)$ and $\mu_X = np = 2$.

Distributions of Discrete Random Variables

The Poisson Distribution

Example 7: A mass contains 10,000 atoms of a radioactive substance. The probability that a given atom will decay in a one-minute time period is 0.0002. Let X represent the number of atoms that decay in one minute.

Each atom can be considered as a Bernoulli trial, where decay of an atom represents 'success'. Since, X denotes the number of successes in 10,000 independent Bernoulli trials, $X \sim \mathrm{Bin}(10000, 0.0002)$ and $\mu_X = np = 2$.

Another mass contains 5,000 atoms, and each of these atoms has probability 0.0004 of decaying in a one-minute time interval.

Distributions of Discrete Random Variables
The Poisson Distribution

Example 7: A mass contains 10,000 atoms of a radioactive substance. The probability that a given atom will decay in a one-minute time period is 0.0002. Let X represent the number of atoms that decay in one minute.

Each atom can be considered as a Bernoulli trial, where decay of an atom represents 'success'. Since, X denotes the number of successes in 10,000 independent Bernoulli trials, $X \sim \mathrm{Bin}(10000, 0.0002)$ and $\mu_X = np = 2$.

Another mass contains 5,000 atoms, and each of these atoms has probability 0.0004 of decaying in a one-minute time interval.

Similarly, considering Y as a random variable denoting the number of atoms that decay in one minute, $Y \sim \mathrm{Bin}(5000, 0.0004)$ and $\mu_X = np = 2$.

Example 7 (Continued): Assuming that we want to compute the probability that exactly six atoms decay in one minute for each of these masses. Using the binomial probability mass function, we have

$$P(X = 6) = \frac{10000!}{6!9994!}(0.0002)^6(0.9998)^{9994} = 0.01202378737$$

$$P(Y = 6) = \frac{5000!}{6!4994!}(0.0004)^6(0.9996)^{4994} = 0.01201777042$$

Distributions of Discrete Random Variables

The Poisson Distribution

Example 7 (Continued): Assuming that we want to compute the probability that exactly six atoms decay in one minute for each of these masses. Using the binomial probability mass function, we have

$$P(X = 6) = \frac{10000!}{6!9994!}(0.0002)^6(0.9998)^{9994} = 0.01202378737$$

$$P(Y = 6) = \frac{5000!}{6!4994!}(0.0004)^6(0.9996)^{4994} = 0.01201777042$$

Therefore, these probabilities are almost the same. When n is large and p is small, pmf depends almost entirely on the mean $\lambda = np$ and we can approximate the binomial distribution with the Poisson distribution where

$$P(X = 6) \approx P(Y = 6) \approx \frac{(e^{-\lambda})(\lambda^6)}{6!} = \frac{(e^{-2})(2^6)}{6!} = 0.01202980295.$$

Distributions of Discrete Random Variables

The Poisson Distribution

The Poisson Distribution ($X \sim \text{Poisson}(\lambda)$)

Consider an interval of real numbers, where an event can occur at random throughout the interval. If the interval is split into small subintervals such that

1. probability of more than one occurrence of the event in a subinterval is zero,
2. probability of one occurrence of the event in a subinterval is the same for all subintervals and proportional to the length of the subinterval, and
3. the occurrence count in each subinterval is independent of the other subintervals,

the random experiment is called a **Poisson process**. The pmf for the **Poisson random variable**, X with parameter λ is given by

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad \text{for} \qquad x = 0, 1, 2\ldots,$$

$\mu_X = E(X) = \lambda$ and $\sigma_X^2 = V(X) = \lambda$.

Distributions of Discrete Random Variables

The Poisson Distribution

Example 8: The number of queries arriving in t seconds at a call center, X, is a Poisson random variable with parameter $\lambda = \alpha t$ where α is the average arrival rate in queries/second. Assume that the arrival rate is 5 queries per minute. Find the probability of the following events: (a) more than 3 queries in 4 seconds; (b) fewer than 5 queries in 2 minutes.

Distributions of Discrete Random Variables

The Poisson Distribution

Example 8: The number of queries arriving in t seconds at a call center, X, is a Poisson random variable with parameter $\lambda = \alpha t$ where α is the average arrival rate in queries/second. Assume that the arrival rate is 5 queries per minute. Find the probability of the following events: (a) more than 3 queries in 4 seconds; (b) fewer than 5 queries in 2 minutes.

Solution (a):

The arrival rate is $\alpha = 5/60 = 1/12$ (queries/sec). Considering the time interval of 10 seconds, $\lambda = 4(\frac{1}{12}) = \frac{1}{3}$. Thus,

$$P(X > 3) = 1 - P(x \leq 3) = 1 - \sum_{x=0}^{3} \frac{(e^{-\frac{1}{3}})(\frac{1}{3})^x}{x!} = 3.95 \times 10^{-4}$$

Solution (b) is left as an exercise.

Distributions of Discrete Random Variables

The Poisson Distribution

Example 9: An optical communication system transmits information at a rate of 1 gigabits/second. The probability of a bit error in the optical communication system is 10^{-9}. Find the probability of five or more errors in 1 second.

Distributions of Discrete Random Variables

The Poisson Distribution

Example 9: An optical communication system transmits information at a rate of 1 gigabits/second. The probability of a bit error in the optical communication system is 10^{-9}. Find the probability of five or more errors in 1 second.

Solution: Transmission of each bit is a Bernoulli trial with a "success" corresponding to a bit error in transmission. The probability of x errors in transmission of $n = 10^9$ bits per second can be calculated by the binomial probability with $n = 10^9$ and $p = 10^{-9}$ or alternatively it can be approximated by Poisson probability where $\lambda = np = 1$. Hence,

$$P(X \geq 5) = 1 - P(x \leq 5) = 1 - \sum_{x=0}^{5} \frac{(e^{-\lambda})(\lambda)^x}{x!} = 1 - \sum_{x=0}^{5} \frac{(e^{-1})(1)^x}{x!} = 0.00366$$

Distributions of Discrete Random Variables
Example

Example 9: The probability that a wafer contains a large particle of contamination is 0.01. If it is assumed that the wafers are independent, what is the probability that exactly 100 wafers need to be analyzed before a large particle is detected?

Distributions of Discrete Random Variables

Example

Example 9: The probability that a wafer contains a large particle of contamination is 0.01. If it is assumed that the wafers are independent, what is the probability that exactly 100 wafers need to be analyzed before a large particle is detected?

Solution: Let X represent the number of samples analyzed until a large particle is detected. Then $X \sim \mathrm{Geom}(p)$, and since $p = 0.01$,

$$P(X = 100) = (1 - 0.01)^{99}(0.01) = 0.0037$$

Distributions of Discrete Random Variables

Example

Example 10: A batch of parts contains 100 parts from a local supplier of tubing and 200 parts from a supplier of tubing in another country. If 4 parts are selected randomly and without replacement, (a) what is the probability they are all from the local supplier? (b) What is the probability that at least one part in the sample is from the local supplier?

Distributions of Discrete Random Variables

Example

Example 10: A batch of parts contains 100 parts from a local supplier of tubing and 200 parts from a supplier of tubing in another country. If 4 parts are selected randomly and without replacement, (a) what is the probability they are all from the local supplier? (b) What is the probability that at least one part in the sample is from the local supplier?

Solution (a): Let X be equal to the number of parts in the sample from the local supplier. Then X has hypergeometric distribution and

$$P(X = 4) = \frac{\binom{100}{4}\binom{200}{0}}{\binom{300}{4}} = 0.0119$$

Distributions of Discrete Random Variables

Example

Example 10: A batch of parts contains 100 parts from a local supplier of tubing and 200 parts from a supplier of tubing in another country. If 4 parts are selected randomly and without replacement, (a) what is the probability they are all from the local supplier? (b) What is the probability that at least one part in the sample is from the local supplier?

Solution (a): Let X be equal to the number of parts in the sample from the local supplier. Then X has hypergeometric distribution and

$$P(X = 4) = \frac{\binom{100}{4}\binom{200}{0}}{\binom{300}{4}} = 0.0119$$

Solution (b): $P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{\binom{100}{0}\binom{200}{4}}{\binom{300}{4}} = 0.804$

Distributions of Discrete Random Variables

Example

Example 11: A research team has developed a face recognition device to match photos in a database. From laboratory tests, the recognition accuracy is 92% and trials are assumed to be independent. (a) If the research team continues to run laboratory tests, what is the mean number of trials until failure? (b) What is the probability that the first failure occurs on the tenth trial? (c) To improve the recognition algorithm, a chief engineer decides to collect 10 failures. How many trials are expected to be needed?

Distributions of Discrete Random Variables
Example

Example 11: A research team has developed a face recognition device to match photos in a database. From laboratory tests, the recognition accuracy is 92% and trials are assumed to be independent. (a) If the research team continues to run laboratory tests, what is the mean number of trials until failure? (b) What is the probability that the first failure occurs on the tenth trial? (c) To improve the recognition algorithm, a chief engineer decides to collect 10 failures. How many trials are expected to be needed?

Solution (a): Let X denote the number of trials until failure. Then X is a geometric random variable with $p = 1 - 0.92 = 0.08$. Thus, $E(X) = 1/p = 12.5$.

Distributions of Discrete Random Variables

Example

Example 11: A research team has developed a face recognition device to match photos in a database. From laboratory tests, the recognition accuracy is 92% and trials are assumed to be independent. (a) If the research team continues to run laboratory tests, what is the mean number of trials until failure? (b) What is the probability that the first failure occurs on the tenth trial? (c) To improve the recognition algorithm, a chief engineer decides to collect 10 failures. How many trials are expected to be needed?

Solution (a): Let X denote the number of trials until failure. Then X is a geometric random variable with $p = 1 - 0.92 = 0.08$. Thus, $E(X) = 1/p = 12.5$.

Solution (b): X denotes the number of tests until the first failure, and X follows a geometric distribution. Thus, $P(X = 10) = (0.92)^9(0.08) = 0.0378$.

Distributions of Discrete Random Variables

Example

Example 11: A research team has developed a face recognition device to match photos in a database. From laboratory tests, the recognition accuracy is 92% and trials are assumed to be independent. (a) If the research team continues to run laboratory tests, what is the mean number of trials until failure? (b) What is the probability that the first failure occurs on the tenth trial? (c) To improve the recognition algorithm, a chief engineer decides to collect 10 failures. How many trials are expected to be needed?

Solution (c): Let Y denote the number of trials needed to collect 10 failures. Then Y is a negative binomial random variable with $p = 0.08$. Thus,
$E(Y) = 10/0.08 = 125$.

Distributions of Discrete Random Variables

Example

Example 12: A six-sided die is tossed 12 times. What is the probability of obtaining a sequence in which each face appears exactly twice?

Distributions of Discrete Random Variables
Example

Example 12: A six-sided die is tossed 12 times. What is the probability of obtaining a sequence in which each face appears exactly twice?

Solution: The result of the trial can be split into 6 classes, represented by random variables $X_1, X_2, X_3, X_4, X_5, X_6$. These random variables have a multinomial distribution, and thus,

$$P(X_1 = 2, X_2 = 2, X_3 = 2, X_4 = 2, X_5 = 2, X_6 = 2) = \frac{12!}{2!2!2!2!2!2!}((\frac{1}{6})^2)^6 \approx 3.4\times 10^{-3}$$