STAT 509: Statistics for Engineers

Chapter 3: Discrete Random Variables and Probability Distributions

Dr. Dewei Wang
Associate Professor
Department of Statistics
University of South Carolina
deweiwang@stat.sc.edu

Chapter 3: Discrete Random Variables and Probability Distributions

Learning Objectives:

1. Understand random variables
2. Determine probabilities from probability mass functions and the reverse
3. Determine probabilities and probability mass functions from cumulative distribution functions and the reverse.
4. Calculate means and variances for discrete random variables.
5. Understand the assumptions for discrete probability distributions.
6. Select an appropriate discrete probability distribution to calculate probabilities.
7. Calculate probabilities, means and variances for discrete probability distributions.

Random Variable and its Notation

A variable that associates a number with the outcome of a random experiment is called a **random variable**; e.g., flip a coin,

$$X = \begin{cases} 1 & \text{head} \\ 0 & \text{tail} \end{cases}$$

A random variable is a **function** that assigns a real number to each outcome in the sample space of a random experiment; e.g., the sample space of flipping a coin is $\mathcal{S} = \{head, tail\}$

$$X : \mathcal{S} \mapsto \{0, 1\} : X(head) = 1 \quad X(tail) = 0.$$

A random variable is denoted by an uppercase letter such as X. After the experiment is conducted, the measured value of the random variable is denoted by a lowercase letter such as $x = 70$ milliamperes.

Discrete and Continuous Random Variables

A **discrete random variable** is a random variable with a finite or countably infinite range. Its values are obtained by counting.

- Number of scratches on a surface
- Proportion of defective parts among 100 tested
- Number of transmitted bits received in error
- Number of common stock shares traded per day

A **continuous random variable** is a random variable with an interval (either finite or infinite) of real numbers for its range. Its values are obtained by measuring.

- Electrical current and voltage
- Physical measurements, e.g., length, weight, time, temperature, pressure

Discrete Random Variables

Many physical systems can be modeled by the same or similar random experiments and random variables. The distribution of the random variables involved in each of these common systems can be analyzed. We will learn many classical random variables that can be used in different applications and examples.

We start with **discrete random variables**.

Notation

Let X be a discrete random variable. Denote its possible values by $x_1, x_2, \ldots, x_n, \ldots$ (could be infinite).

We would like to know the probability $P(X = x_i)$ for an x_i or $P(X \leq x)$ for any x.

Probability Distributions and Probability Mass Functions

The time to recharge the flash is tested in three cell-phone cameras. The probability that a camera passes the test is 0.8, and the cameras perform independently. The sample space for this experiment and associated probabilities are:

TABLE 3.1 **Camera Flash Tests**

Camera 1	Camera 2	Camera 3	Probability	X
Pass	Pass	Pass	0.512	3
Fail	Pass	Pass	0.128	2
Pass	Fail	Pass	0.128	2
Fail	Fail	Pass	0.032	1
Pass	Pass	Fail	0.128	2
Fail	Pass	Fail	0.032	1
Pass	Fail	Fail	0.032	1
Fail	Fail	Fail	0.008	0

where the random variable X denotes the number of cameras that pass the test. Then the possible value

Probability Mass Function

For a discrete random variable X with possible values $x_1, \ldots, x_n$ (if finite) or $x_1, x_2, \ldots, x_n, \ldots$ (if infinite), its **probability mass function** (pmf) is a function such that:

1. $f(x_i) = P(X = x_i)$
2. $f(x_i) \geq 0$
3. $\sum_{i=1}^{n} f(x_i) = 1$ (if finite) or $\sum_{i=1}^{\infty} f(x_i) = 1$ (if infinite).

We often call $f(x)$ as the probability density function (pdf) of X as well, though rigorously, pdf is for a continuous random variable.

Example: infinite possible values

Let the random variable X denote the number of wafers that need to be analyzed to detect a large particle of contamination. Assume that the probability that a wafer contains a large particle is 0.01, and that the wafers are independent. Determine the probability distribution of X.

Answer: Let p denote a wafer in which a large particle is present & let a denote a wafer in which it is absent.

- The sample space is $S = \{p, ap, aap, aaap, aaaap, \dots\}$
- The associated range of X is $x_1, x_2, x_3, x_4, x_5, \dots$ where $x_i = i$.

We then have

$$P(X = 1) = 0.01, \quad P(X = 2) = 0.99 \cdot 0.01, \quad P(X = 3) = 0.99^2 \cdot 0.01,$$

so and so on; i.e., for $i \geq 1$,

$$P(X = x_i) = P(X = i) = 0.99^{i-1} \cdot 0.01.$$

Cumulative Distribution Function

The **cumulative distribution function** (cdf), is the probability that a random variable X will be found at a value less than or equal to x. Symbolically, the cdf is

$$F(x) = P(X \leq x).$$

For a discrete random variable X, $F(x)$ satisfies the following properties:

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

Note that pmf f is defined on all the possible values of X while cdf F is defined on the entire real line.

Example: cdf

Suppose X is a discrete random variable with pmf being

$$f(x) = \begin{cases} 0.2 & x = -2 \\ 0.5 & x = 0 \\ c & x = 2. \end{cases}$$

Determine and draw the cdf of X.

Answer: the definition of $f(x)$ suggests that X has only three possible values $\{-2, 0, 2\}$. Known from the definition of a pmf; i.e., $\sum_{i=1}^{n} f(x_i) = 1$, we have

$$1 = f(-2) + f(0) + f(2) = 0.2 + 0.5 + c$$

which implies $c = 0.3$.

Example: cdf continued

Now we find the cdf $F(x)$ of X. Because $F(x)$ is defined on the entire real line, and X only has three possible values which partition the entire real line into four regions:

$$(-\infty, \underbrace{-2)}_{x_1} \cup \underbrace{[-2, 0)}_{x_2} \cup \underbrace{[0, 2)}_{x_3} \cup [2, \infty).$$

- For $x \in (-\infty, -2)$, $F(x) = P(X \leq x) = 0$
- For $x \in [-2, 0)$, $F(x) = P(X \leq x) = P(X = -2) = f(-2) = 0.2$
- For $x \in [0, 2)$, $F(x) = P(X \leq x) = P(X = -2) + P(X = 0) = f(-2) + f(0) = 0.7$
- For $x \in [-2, 0)$, $F(x) = P(X \leq x) = P(X = -2) + P(X = 0) + P(X = 2) = f(-2) + f(0) + f(2) = 1$

Example: cdf continued

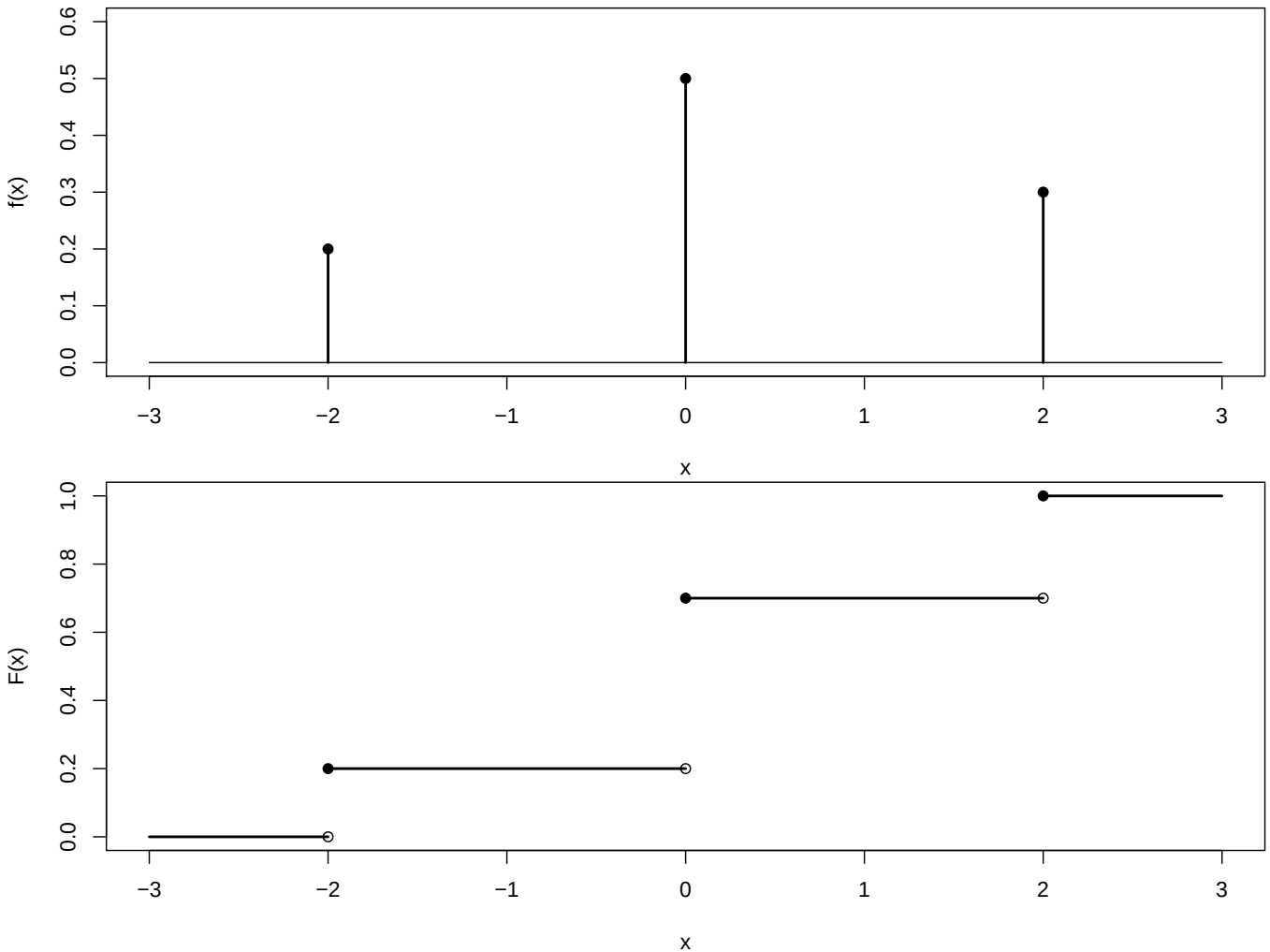
Now we have

$$F(x) = \begin{cases} 0 & x < -2 \\ 0.2 & -2 \leq x < 0 \\ 0.7 & 0 \leq x < 2 \\ 1 & 2 \leq x. \end{cases}$$

```
#StatEngine
x=c(-2,0,2);fx=c(0.2,0.5,0.3)
discrete.plotcdf(x,fx)
# or
discrete.summary(x,fx,plotpdf=FALSE,plotcdf=TRUE)
```

Example: pmf/cdf plot

```
#StatEngine
x=c(-2,0,2);fx=c(0.2,0.5,0.3)
discrete.summary(x,fx)
```



Mean and Variance of a Discrete Random Variable

Two numbers are often used to summarize a probability distribution for a random variable X.

- The **mean** is a measure of the center or middle of the probability distribution.
- The **variance** is a measure of the dispersion, or variability in the distribution.

These two measures **do not** uniquely identify a probability distribution. That is, two different distributions can have the same mean and variance. Still, these measures are simple, useful summaries of the probability distribution of X.

Mean and Variance of a Discrete Random Variable

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

$$\mu = E(X) = \sum_i x_i f(x_i).$$

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

$$\sigma^2 = V(X) = E(X-\mu)^2 = \sum_i (x_i-\mu)^2 f(x_i) = \sum_i x_i^2 f(x_i) - \mu^2.$$

The **standard deviation** of X is

$$\sigma = \sqrt{\sigma^2}$$

```
# StatEngine
discrete.summary(x,fx)
```

Mean and Variance of a Discrete Random Variable

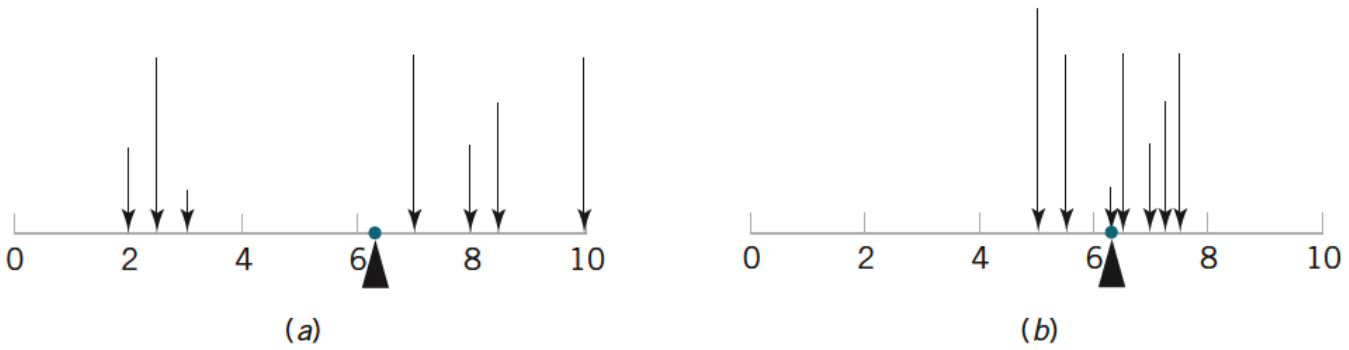


FIGURE 3.4

A probability distribution can be viewed as a loading with the mean equal to the balance point. Parts (a) and (b) illustrate equal means, but part (a) illustrates a larger variance.

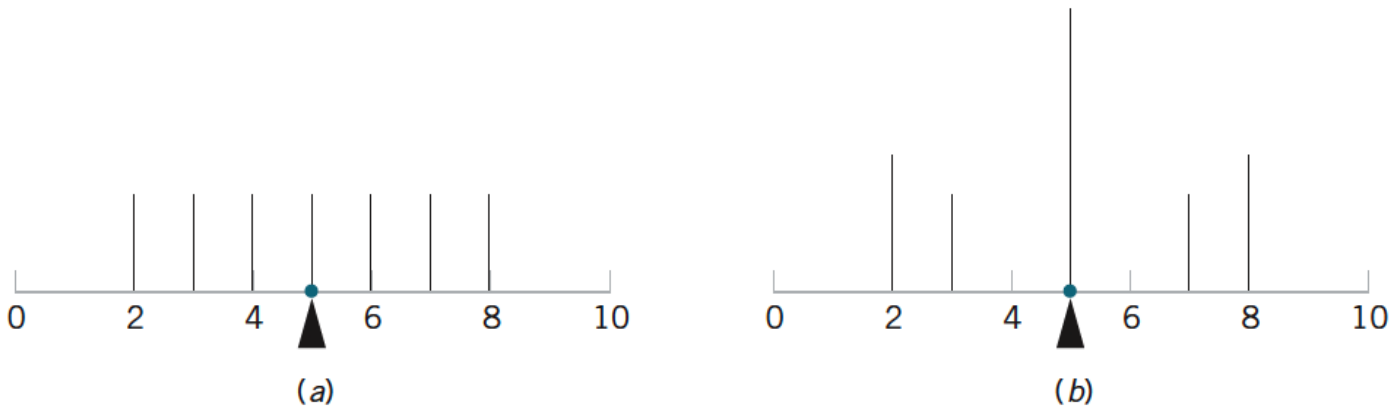


FIGURE 3.5

The probability distributions illustrated in parts (a) and (b) differ even though they have equal means and equal variances.

Example: Mean and Variance

Suppose X is a discrete random variable with pmf being

$$f(x) = \begin{cases} 0.2 & x = -2 \\ 0.5 & x = 0 \\ 0.3 & x = 2. \end{cases}$$

```
#StatEngine
x=c(-2,0,2);fx=c(0.2,0.5,0.3);
discrete.summary(x,fx)
$mean
[1] 0.2
$variance
[1] 1.96
$standard.deviation
[1] 1.4
```

Expected Value of a Function of a Discrete X

If X is a discrete random variable with pmf $f(x)$ on possible values x_i's, then for any real function h,

$$E[h(X)] = \sum_i h(x_i) f(x_i).$$

Note that $V(X) = E(X - \mu)^2 = E[h(X)]$ where $h(x) = (x - \mu)^2$.

If $h(x)$ is linear in x; i.e.,

$$h(x) = ax + b, \text{ where } a, b \text{ are two constants.}$$

Then

- $E[h(X)] = E(aX + b) = aE(X) + b$
- $V[h(X)] = V(aX + b) = a^2 V(X)$.

Example

There is a chance that a bit transmitted through a digital transmission channel is received in error. Let X equal the number of bits in error in the next four bits transmitted. The possible values for X are 0, 1, 2, 3, 4. Based on a model for the errors that is presented in the following section, probabilities for these values will be determined. Suppose that the probabilities are

$$P(X = 0) = 0.6561 \; P(X = 1) = 0.2916$$
$$P(X = 2) = 0.0486 \; P(X = 3) = 0.0036$$
$$P(X = 4) = 0.0001$$

Find $E(X)$, $V(X)$, $E(2X + 4)$, $V(2X + 4)$, $E\{\sin(e^X)\}$ and $E(X^2)$

Example (continued)

Answer: $E(X) = 0.4$ and $V(X) = 0.36$ can be found via

```
x=c(0,1,2,3,4);fx=c(0.6561,0.2916,0.0486,0.0036,0.0001);
discrete.summary(x,fx)
```

$E(2X + 4) = 2E(X) + 4 = 4.8$ and $V(2X + 4) = 4V(X) = 1.44$.

To find $E\{\sin(e^X)\}$, you should use R. Herein, $h(x) = \sin(e^x)$. Define h in R:

```
h=function(x){sin(exp(x))};
```

Then $E\{sin(e^X)\}$ is

```
sum(h(x)*fx)
[1] 0.7186215
```

For $E(X^2)$, you can simply use

```
sum(x^2*fx)
[1] 0.52
```

StatEngine: user-defined discrete distribution

```
discrete.summary(x,fx,plotpdf=c("TRUE","FALSE"),
        plotcdf=c("TRUE","FALSE"))

discrete.prob(x,fx,lb)

discrete.prob(x,fx,lb,ub,
        inclusive=c("none","left","right","both"))
```

Example: Go back to the bit transmission example. The discrete distribution can be defined by

```
x=c(0,1,2,3,4);
fx=c(0.6561,0.2916,0.0486,0.0036,0.0001);
# Some calculation:
discrete.summary(x,fx) # Mean, Variance, std
discrete.prob(x,fx,2) # P(X=2)
discrete.prob(x,fx,3,4,"right") # P(3<X<=4)
```

Classical Discrete Distributions

In the rest of this chapter, we will learn several commonly used discrete distributions:

- Discrete Uniform Distribution
- Binomial Distribution
- Geometric Distribution
- Negative Binomial Distribution
- Hypergeometric Distribution
- Poisson Distribution

You must know:

1. Definition (the right tool for the right job)
2. Probability calculation
3. Mean and variance calculation

Discrete Uniform Distribution

The simplest discrete random variable is one that assumes only a finite number of possible values, each with equal probability.

Discrete Uniform Distribution

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

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

Suppose the range of X equals the **consecutive integers** $a, a+1, a+2, \ldots, b$ for $a \leq b$. The range of X contains $b - a + 1$ values each with probability $1/(b - a + 1)$. Then the mean and variance X are

$$\mu = E(X) = \frac{b+a}{2} \text{ and } \sigma^2 = V(X) = \frac{(b-a+1)^2 - 1}{12}.$$

Example

Suppose that the *discrete uniform* random variable Y has range $5, 10, 15, \ldots, 30$. Find $P(Y \leq 16)$, $E(Y)$ and $V(Y)$.
Answer: If $Y \leq 16$, Y can only be $5, 10$ or 15. Thus $P(Y \leq 16) = P(Y = 5) + P(Y = 10) + P(Y = 15) = 1/6 + 1/6 + 1/6 = 0.5$.

```
range=seq(5,30,by=5);
duniform.prob(range,-Inf,16,inclusive="right")
```

To find $E(Y)$ and $V(Y)$, you can use

```
range=seq(5,30,by=5);duniform.summary(range)
```

Or you could do it in this way:
Obviously, $Y = 5X$ where X is a discrete uniform random variable with range $1, 2, \ldots, 6$. Known from the formula that $a = 1$, $b = 6$,

$$E(X) = \frac{a+b}{2} = 3.5 \text{ and } V(X) = \frac{(6-1+1)^2 - 1}{12} = \frac{35}{12}.$$

Then using $Y = 5X$, we have

$$E(Y) = 5E(X) = 17.5 \text{ and } V(Y) = 25V(X) = 72.92.$$

StatEngine: discrete uniform distribution

```
duniform.summary(range,plotpdf=c("TRUE","FALSE"),
        plotcdf=c("TRUE","FALSE"))

duniform.prob(range,lb)

duniform.prob(range,lb,ub,
        inclusive=c("none","left","right","both"))
```

Example: Go back to the previous example. The uniform distribution of Y can be defined by

```
range=seq(5,30,by=5);
# Some calculation:
duniform.summary(range) # Mean, Variance, std
duniform.prob(range,2) # P(X=2)
duniform.prob(range,-Inf,16,"right") # P(X<=16)
```

Bernoulli Distribution (Bernoulli Trials) $X \sim \text{Bernoulli}(p)$

The next three distributions (Binomial, Geometric, Negative Binomial) are from (a series of independent and identical) Bernoulli trials.

Definition

A trial with only two possible outcomes (often labeled as "success" and "failure") called a **Bernoulli trial**. The associated distribution is called a **Bernoulli distribution**.

A random variable X is said to follow a Bernoulli distribution with probability of success p, denoted by $X \sim \text{Bernoulli}(p)$, if X only takes two values, 0 and 1 and

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

E.g., flip a coin; a part is defect or not; air contains contamination or not; the next birth is a boy or a girl; medication is effective or not.

Binomial, Geometric, and Negative Binomial

Provided with a Bernoulli trial where the probability of success is p.

- Suppose we independently and identically repeat the trial for n times, let X be the **number of successes** from the n trials. Then X follows a Binomial distribution.
- Suppose we independently and identically repeat the trial, let X be the number of trial to observe the **first** success. Then X follows a Geometric distribution.
- Suppose we independently and identically repeat the trial, let X be the number of trial to observe the r-**th** success. Then X follows a Negative Binomial distribution.

Binomial Distribution (Definition) $X \sim \text{Binomial}(n, p)$

A random experiment consists of n Bernoulli trials such that

(1) The trials are independent.

(2) Each trial results in only two possible outcomes, labeled as "success" and "failure."

(3) The probability of a success in each trial, denoted as p, remains constant.

The random variable X that equals the number of trails that result in a success is a **binomial random variable** with parameters $0 < p < 1$ and n, denoted by $X \sim \text{Binomial}(n, p)$. The probability mass function of X is

$$f(x) = C_x^n p^x (1-p)^{n-x} \quad \text{for } x = 0, 1, \ldots, n.$$

Its mean and variances are

$$\mu = E(X) = np \quad \text{and} \quad \sigma^2 = V(X) = np(1-p).$$

Binomial Distribution (Examples)

1. Flip a coin 10 times. Let $X = \#$ of heads obtained.
2. A worn machine tool produces 1% defective parts. $X = \#$ of defective parts in the next 25 parts produced.
3. Each sample of air has a 10% chance of containing a particular rare molecule. $X = \#$ of air samples that contain the rare molecule in the next 18 samples analyzed.
4. Of all bits transmitted through a digital transmission channel, 10% are received in error. $X = \#$ of bits in error in the next five bits transmitted.
5. A multiple-choice test contains 10 questions, each with four choices, and you guess at each question. $X = \#$ of questions answered correctly.
6. In the next 20 births at a hospital, $X = \#$ of female births.
7. Of all patients suffering a particular illness, 35% experience improvement from a medication. In the next 100 patients administered the medication, $X = \#$ of patients who experience improvement.

Binomial Distribution (Example)

Each sample of water has a 10% chance of containing a particular organic pollutant. Assume that the samples are independent with regard to the presence of the pollutant. Find the probability that in the next 18 samples, exactly 2 contain the pollutant.

Answer: Let $X =$ the number of samples that contain the pollutant in the next 18 samples analyzed. We know that $X \sim \text{Binomial}(n = 18, p = 0.1)$. Thus

$$P(X = 2) = f(2) = C_2^{18} 0.1^2 (1 - 0.1)^{16} = 0.284.$$

In addition, $P(X > 4) = \sum_{x=5}^{18} f(x) = \sum_{x=5}^{18} C_x^{18} 0.1^x (1-0.1)^{18-x}$ or $P(X > 4) = 1 - P(X \leq 4) = 1 - \sum_{x=0}^{4} C_x^{18} 0.1^x (1 - 0.1)^{18-x}$. Also, $P(3 \leq X < 7) = \sum_{x=3}^{6} C_x^{18} 0.1^x (1 - 0.1)^{18-x}$ or

$$P(3 \leq X < 7) = P(X \leq 6) - P(X \leq 2). \textit{Why?}$$

The mean and variance of X are $\mu = np = 1.8$ and $\sigma^2 = np(1-p) = 1.62$.

StatEngine: binomial distribution

```
binomial.summary(n,p,plotpdf=c("TRUE","FALSE"),
        plotcdf=c("TRUE","FALSE"))

binomial.prob(n,p,lb)

binomial.prob(n,p,lb,ub,
        inclusive=c("none","left","right","both"))
```

Example: Go back to the previous example. The binomial distribution can be defined by

```
n=18;p=0.1;
# Some calculation:
binomial.summary(n,p) # Mean, Variance, std
binomial.prob(n,p,2) # P(X=2)
binomial.prob(n,p,3,7,"left") # P(3<=X<7)
```

Geometric Distribution (Definition) $X \sim \text{Geometric}(p)$

In a series of Bernoulli trials (independent trials with constant probability p of a success), the random variable X that equals the number of trials until the **first** success is a **geometric random variable** with parameter $0 < p < 1$, denoted by $X \sim \text{Geometric}(p)$. The probability mass function of X is

$$f(x) = (1-p)^{x-1}p \quad \text{for } x = 1, 2 \ldots.$$

Its cdf is

$$F(x) = 1 - (1-p)^{x} \quad \text{for } x = 1, 2 \ldots.$$

Its mean and variances are

$$\mu = E(X) = 1/p \quad \text{and} \quad \sigma^2 = V(X) = (1-p)/p^2.$$

Geometric Distribution (Example)

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 125 wafers need to be analyzed before a large particle is detected?

Answer: Let X denote the number of samples analyzed until a large particle is detected. Then $X \sim \text{Geometric}(p = 0.01)$. The requested probability is

$$P(X = 125) = 0.99^{124}0.01 = 0.0029.$$

In addition, the $\mu = 1/p = 100$ and $\sigma^2 = (1 - p)/p^2 = 9900$.

StatEngine: geometric distribution

```
geometric.summary(p,plotpdf=c("TRUE","FALSE"),
        plotcdf=c("TRUE","FALSE"))

geometric.prob(p,lb)

geometric.prob(p,lb,ub,
        inclusive=c("none","left","right","both"))
```

Example: Go back to the previous example. The geometric distribution can be defined by

```
p=0.01;
# Some calculation:
geometric.summary(p) # Mean, Variance, std
geometric.prob(p,125) # P(X=125)
geometric.prob(p,3,Inf,"none") # P(3<X)
```

Geometric Distribution (Lack of Memory)

A geometric random variable has been defined as the number of trials until the first success. However, because the trials are independent, the count of the number of trials until the next success can be started at any trial without changing the probability distribution of the random variable.

For example, if 100 bits are transmitted, the probability that the first error, after bit 100, occurs on bit 106 is the probability that the next six outcomes are $OOOOOE$ (O means okay, E means error). This probability is $(0.9)^5(0.1) = 0.059$, which is identical to the probability that the initial error occurs on bit 6.

The implication of using a geometric model is that the system presumably does not wear out. The probability of an error remains constant for all transmissions. In this sense, the geometric distribution is said to lack any memory. The **lack of memory property** is discussed again in the context of an exponential random variable in a later chapter.

Negative Binomial Distribution (Definition)
$X \sim \text{NegBinom}(r, p)$

In a series of Bernoulli trials (independent trials with constant probability p of a success), the random variable X that equals the number of trials until the r-**th** success is a **negative binommial random variable** with parameters $0 < p < 1$ and r, denoted by $X \sim \text{NegBinom}(r, p)$. The probability mass function of X is

$$f(x) = C_{r-1}^{x-1}(1-p)^{(x-r)}p^r \quad \text{for } x = r, r+1, r+2 \ldots.$$

Its mean and variances are

$$\mu = E(X) = r/p \quad \text{and} \quad \sigma^2 = V(X) = r(1-p)/p^2.$$

Negative Binomial Distribution (Example)

The probability that a camera passes the test is 0.8, and the cameras perform independently. What is the probability that the third failure is obtained in five or fewer tests?

Answer: Let X denote the number of cameras tested until three failures have been obtained. Then $X \sim \text{NegBinom}(r = 3, p = 0.2)$. The requested probability is

$$P(X \leq 5) = \sum_{x=3}^{5} C_2^{x-1}(1-0.2)^{(x-3)}0.2^3 = 0.058.$$

Its mean and variance are $\mu = r/p = 3/0.2 = 15$ and $\sigma^2 = r(1-p)/p^2 = 3(0.8)/0.04 = 60$.

StatEngine: negative binomial distribution

```
negbinom.summary(r,p,plotpdf=c("TRUE","FALSE"),
        plotcdf=c("TRUE","FALSE"))

negbinom.prob(r,p,lb)

negbinom.prob(r,p,lb,ub,
        inclusive=c("none","left","right","both"))
```

Example: Go back to the previous example. The negative binomial distribution can be defined by

```
r=3;p=0.2;
# Some calculation:
negbinom.summary(r,p) # Mean, Variance, std
negbinom.prob(r,p,5) # P(X=5)
negbinom.prob(r,p,-Inf,5,"right") # P(X<=5)
```

Hypergeometric Distribution (Definition)

$X \sim \text{HyperGeo}(N, K, n)$

Example: A day's production of 850 manufactured parts contains 50 parts that do not conform to customer requirements. Two parts are selected at random without replacement from the day's production.

Let A and B denote the events that the first and second parts are nonconforming, respectively. From counting parts in the sample space, $P(B|A) = 49/849$ and $P(A) = 50/850$. Consequently, knowledge that the first part is nonconforming suggests that it is less likely that the second part selected is nonconforming; i.e., A and B are not independent. Let X equal the number of nonconforming parts in the sample of size 2. Then

$$P(X = 0) = \frac{800}{850}\frac{799}{849}, \quad P(X = 2) = \frac{50}{850}\frac{49}{849} = 0.003,$$

$$P(X = 1) = \frac{800}{850}\frac{50}{849} + \frac{50}{850}\frac{800}{849} = 0.111.$$

Hypergeometric Distribution (Definition)

$X \sim \text{HyperGeo}(N, K, n)$

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 $n \leq N$.

The random variable X that equals the number of successes in the sample of size n is a **hypergeometric random variable**, denoted by $X \sim \text{HyperGeo}(N, K, n)$, and its pmf is

$$f(x) = \frac{C_x^K C_{n-x}^{N-K}}{C_n^N}, \quad \text{from } x = \max\{0, n + K - N\} \text{ to } \min\{K, n\}.$$

The mean and variance are (similar to Binomial distribution)

$$\mu = E(X) = np \quad \text{and} \quad \sigma^2 = V(X) = np(1-p)\left(\frac{N-n}{N-1}\right),$$

where $p = K/N$ and $\frac{N-n}{N-1}$ is the finite Population Correction Factor.

Hypergeometric Distribution (Example)

A batch of parts contains 100 from a local supplier of circuit boards and 200 from a supplier in the next state. If four parts are selected randomly and without replacement, what is the probability they are all from the local supplier?

Answer: Let X equal the number of parts in the sample from the local supplier. Then $X \sim \text{HyperGeo}(N = 300, K = 100, n = 4)$. The requested probability is

$$P(X = 4) = \frac{C_4^{100} C_0^{200}}{C_4^{300}} = 0.0119.$$

The mean and variance are $\mu = nK/N = 4/3$ and $\sigma^2 = n(K/N)(1-K/N)(N-n)/(N-1) = 4(100/300)(200/300)296/299 = 0.88$

StatEngine: hypergeometric distribution

```
hypergeo.summary(N,K,n,plotpdf=c("TRUE","FALSE"),
        plotcdf=c("TRUE","FALSE"))

hypergeo.prob(N,K,n,lb)

hypergeo.prob(N,K,n,lb,ub,
        inclusive=c("none","left","right","both"))
```

Example: Go back to the previous example. The hypergeometric distribution can be defined by

```
N=300;K=100;n=4;
# Some calculation:
hypergeo.summary(N,K,n) # Mean, Variance, std
hypergeo.prob(N,K,n,4) # P(X=4)
hypergeo.prob(N,K,n,1,2,"right") # P(1<X<=2)
```

Poisson Distribution (Definition) $X \sim \text{Poisson}(\lambda, L)$

A Poisson random variable is the number of event occurs in an interval of length T (time, area, volume), where the occurrence of the events follows a Poisson process.

Example

1. the number of flaws in a length of 10 millimeters of wire
2. the number of customers who enter a bank in an hour
3. the number of stars in a given volume of space

A Poisson Process is a random experiment with properties:

In general, consider subintervals of small length $\Delta t \approx 0$:

1. The probability of more than one event in a subinterval tends to zero.
2. The probability of one event in a subinterval tends to Δt.
3. The event in each subinterval is independent of other subintervals.

Poisson Distribution (Definition) $X \sim \text{Poisson}(\lambda, L)$

The random variable X that equals the number of events in an interval of length L, where the occurrence follows a Poisson process with a mean rate λ (events per unit length). Then X is a **Poisson random variable** with parameter $\lambda > 0$ and $L > 0$, denoted by $X \sim \text{Poisson}(\lambda, L)$, and the pmf is

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

Its mean and variance are $\mu = E(X) = \lambda L$ and $\sigma^2 = V(X) = \lambda L s$.

Poisson Distribution (Example)

Flaws occur at random along the length of a thin copper wire. Suppose that the number of flaws follows a Poisson distribution with a mean of 2.3 flaws per millimeter (Thus, we know $\lambda = 2.3$ flaws per millimeter).

Determine the probability of 10 flaws in 5 millimeters of wire.

Answer: Let X denote the number of flaws in 5 millimeters of wire. Then $X \sim \text{Poisson}(\lambda = 2.3, L = 5)$ and $\lambda L = 11.5$.

$$P(X = 10) = e^{-11.5}\frac{11.5^{1}0}{10!} = 0.113.$$

Determine the probability of at least one flaw in 2 millimeters.

Answer: Let X denote the number of flaws in 2 millimeters of wire. Then $X \sim \text{Poisson}(\lambda = 2.3, L = 2)$ and $\lambda L = 4.6$. $P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-4.6} = 0.9899$.

Determine the mean and variance of flaws in 10 millimeters of wire.

Answer: Let X denote the number of flaws in 10 millimeters of wire. Then $X \sim \text{Poisson}(\lambda = 2.3, L = 10)$. Thus, $E(X) = V(X) = \lambda L = 23$

StatEngine: Poisson distribution

```
poisson.summary(lambda,L,plotpdf=c("TRUE","FALSE"),
        plotcdf=c("TRUE","FALSE"))

poisson.prob(lambda,L,lb)

poisson.prob(lambda,L,lb,ub,
        inclusive=c("none","left","right","both"))
```

Example: Go back to the previous example. if $L = 5$, the Poisson distribution can be defined by

```
lambda=2.3;L=5
poisson.prob(lambda,L,10) # P(X=10)
# If L=2
poisson.prob(lambda,2,1,Inf,"left") # P(1<=X)
# If L=10
poisson.summary(lambda,10)
```