PROBABILITY AND STATISTICS

CHAPTER 3: SOME SPECIAL DISTRIBUTIONS.

Dr. Phan Thi Huong

HoChiMinh City University of Technology
Faculty of Applied Science, Department of Applied Mathematics
Email: huongphan@hcmut.edu.vn

HCM city — 2021.

OUTLINE

1. DISCRETE DISTRIBUTIONS

OUTLINE

THE BINOMIAL DISTRIBUTION $B(n,p)$

DEFINITION 1.1

- Bernoulli trial $Z \sim B(p)$:

u	1	0
P(Z=u)	p	q=1-p

THE BINOMIAL DISTRIBUTION $B(n, p)$

DEFINITION 1.1

- Bernoulli trial $Z \sim B(p)$:

u	1	0
P(Z=u)	p	q=1-p

$\mathbb{E}(Z) = p$ and $\mathbb{V}(Z) = pq$.

THE BINOMIAL DISTRIBUTION $B(n,p)$

DEFINITION 1.1

- Bernoulli trial $Z \sim B(p)$:

u	1	0
P(Z=u)	p	q=1-p

  $\mathbb{E}(Z) = p$ and $\mathbb{V}(Z) = pq$.

- Binomial random variable $X \sim B(n,p)$= the number of successes in n Bernoulli trials (the probability of success in each trial is $0 \leq p \leq 1$).

THE BINOMIAL DISTRIBUTION $B(n,p)$

DEFINITION 1.1

- Bernoulli trial $Z \sim B(p)$:

u	1	0
P(Z=u)	p	q=1-p

$\mathbb{E}(Z) = p$ and $\mathbb{V}(Z) = pq$.

- Binomial random variable $X \sim B(n,p)$= the number of successes in n Bernoulli trials (the probability of success in each trial is $0 \le p \le 1$).
- X = sum of n independent Bernoulli random variables.

$$X = Z_1 + Z_2 + \cdots + Z_n,$$

where $z_i \sim B(p)$.

THE BINOMIAL DISTRIBUTION $B(n, p)$

Some examples of binomial random variables:

- The number of defective items among 20 independent items with the defective rate 5%.
- The number of winning tickets among 11 independent lottery tickets with the winning rate 1%.
- The number of patients reporting symptomatic relief with a specific medication with the effective rate 80%.

THE BINOMIAL DISTRIBUTION $B(n,p)$

DEFINITION 1.2

The number of combinations, subsets of k elements that can be selected from n elements, is denoted as $\binom{k}{n}$ or C_n^k and

$$Cn^k = \binom{k}{n} = \frac{n!}{k!(n-k)!}$$

EXAMPLE 1

A printed circuit board has eight different locations in which a component can be placed. If five identical components are to be placed on the board, how many different designs are possible?

THE BINOMIAL DISTRIBUTION $B(n,p)$

PROPOSITION 1.1

1. X takes values in $\Omega = \{0, 1, \ldots, n\}$ such that

$$f(k) = P(X=k) = \binom{n}{k} p^k q^{n-k}.$$

- $k=0$: $\boxed{q}\ \boxed{q}\ \boxed{q}\ \boxed{\cdots}\ \boxed{q} \implies P(X=0) = q^n.$
- $k=n$: $\boxed{p}\ \boxed{p}\ \boxed{p}\ \boxed{\cdots}\ \boxed{p} \implies P(X=n) = p^n.$
- $k=1$: $\boxed{p}\ \boxed{q}\ \boxed{q}\ \boxed{\cdots}\ \boxed{q} \implies P(X=1) = \binom{n}{1} pq^{n-1}.$
- $k=2$: $\boxed{p}\ \boxed{p}\ \boxed{q}\ \boxed{\cdots}\ \boxed{q} \implies P(X=2) = \binom{n}{2} p^2 q^{n-2}.$

2. $\mathbb{E}(X) = np$ and $\mathbb{V}(X) = npq$.

THE BINOMIAL DISTRIBUTION $B(n,p)$

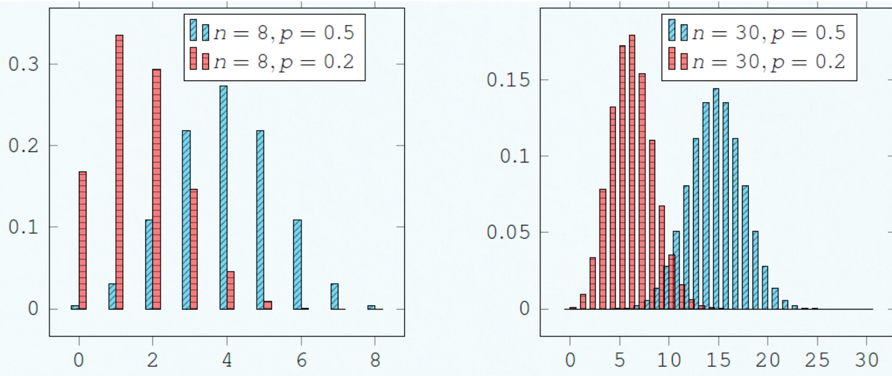


Figure: Pmf of $B(n,p)$

The binomial distribution $B(n, p)$

Example 2

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.

- **A** Find the probability that, in the next 18 samples, exactly 2 contain the pollutant.
- **B** Determine the probability that at least 4 samples contain the pollutant.
- **C** Now determine the probability that $3 \le X < 7$.

THE BINOMIAL DISTRIBUTION $B(n, p)$

EXAMPLE 3

A certain electronic system contains 10 components. Suppose that the probability that each individual component will fail is 0.2 and that the components fail independently of each other. Given that at least one of the components has failed, what is the probability that at least two of the components have failed?

THE HYPERGEOMETRIC DISTRIBUTIONS $H(N; m; n)$

DEFINITION 1.3

- Suppose that there are n draws from a finite population of size N containing m successes without replacement. Let X be the number of successes. Then X is called a hypergeometric random variable or X has a hypergeometric distribution $X \sim H(N, m, n)$.
- X is a sum of n dependent Bernoulli random variables.

$$X = Z_1 + Z_2 + \cdots + Z_n, \quad Z_i = \begin{cases} 1, & \text{if the i-th trial is successful} \\ 0, & \text{otherwise.} \end{cases}$$

THE HYPERGEOMETRIC DISTRIBUTIONS $H(N; m; n)$

PROPOSITION 1.2

Let $X \sim B(n, p)$. Then

THE HYPERGEOMETRIC DISTRIBUTIONS $H(N; m; n)$

PROPOSITION 1.2

Let $X \sim B(n, p)$. Then

1. X has the probability mas function: $f(k) = \dfrac{\binom{k}{m}\binom{n-k}{N-m}}{\binom{n}{N}}$.

The hypergeometric distributions $H(N;m;n)$

Proposition 1.2

Let $X \sim B(n,p)$. Then

1. X has the probability mas function: $f(k) = \frac{\binom{k}{m}\binom{n-k}{N-m}}{\binom{n}{N}}$.
2. $\mathbb{E}(X) = np$ and $\mathbb{V}(X) = npq \cdot \frac{N-n}{N-1}$ with $p = m/N$.

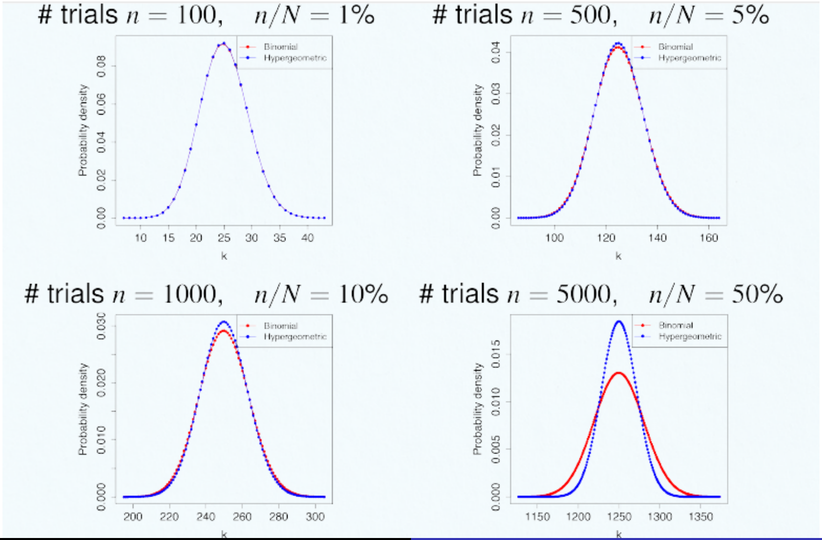
THE HYPERGEOMETRIC DISTRIBUTIONS $H(N;m;n)$

PROPOSITION 1.2

Let $X \sim B(n,p)$. Then

1. X has the probability mas function: $f(k) = \frac{\binom{k}{m}\binom{n-k}{N-m}}{\binom{n}{N}}$.
2. $\mathbb{E}(X) = np$ and $\mathbb{V}(X) = npq \cdot \dfrac{N-n}{N-1}$ with $p = m/N$.
3. The term $\dfrac{N-n}{N-1}$ is called the finite population correction factor. If n is small relative to N ($n << N$), then a binomial distribution can effectively approximate the hypergeometric distribution.

THE HYPERGEOMETRIC DISTRIBUTIONS $H(N; m; n)$



THE HYPERGEOMETRIC DISTRIBUTIONS $H(N; m; n)$

EXAMPLE 4

A batch of parts contains 100 from a local supplier of tubing and 200 from a supplier of tubing in a nearby city. If four 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 two or more parts are from the local supplier?
- C What is the probability that at least one part in the sample is from the local supplier?

The hypergeometric distributions $H(N; m; n)$

Example 5

Suppose that seven balls are selected at random without replacement from a box containing five red balls and ten blue balls. If X denotes the proportion of red balls in the sample, what are the mean and the variance of X?

The Poisson distributions $P(\lambda)$

Example 6

A communication has transmitted n bits. Let X be the number of errors in n transmitted bits. Supposed that the probability of error is a constant p and errors are independent between bits. Let $\lambda = np$, then we have $\mathbb{E}(X) = np = \lambda$ and

$$\mathbb{P}(X=k) = C_n^k p^k (1-p)^{n-k} = C_n^k \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}$$

If the number of transmitted bits is increased and the probability of getting error is decreased such that $\mathbb{E}(X) = \lambda = np$ is a constant, we can prove that

$$\lim_{n\infty} \mathbb{P}(X=k) = e^{-\lambda}\frac{\lambda^k}{k!},\ k = 0,1,2,\ldots$$

The Poisson distributions $P(\lambda)$

Definition 1.4

- A Poisson random variable $X \sim P(\lambda)$ describes the number of events that occur with an unit length where λ is the average number of events per units.

THE POISSON DISTRIBUTIONS $P(\lambda)$

DEFINITION 1.4

- A Poisson random variable $X \sim P(\lambda)$ describes the number of events that occur with an unit length where λ is the average number of events per units.

Some examples:

- The number of misprints on a page of a book
- The number of people in a community who are at least 100 years old.
- The number of people entering a post office on a given day.

Properties of Poisson distributions

PROPERTIES OF POISSON DISTRIBUTIONS

1. The probability mass function of X:

$$f(k) = P(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}, \quad k = 0,\ 1,\ 2,\ 3,\ ,\ldots$$

PROPERTIES OF POISSON DISTRIBUTIONS

1. The probability mass function of X:

$$f(k) = P(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}, \quad k = 0,\ 1,\ 2,\ 3,\ ,\ldots$$

2. The mean and variance of the Poisson model are the same.

$$\boxed{\mathbb{E}(X) = \lambda \text{ and } \mathbb{V}(X) = \lambda.}$$

Otherwise, the Poisson distribution would not be a good model.

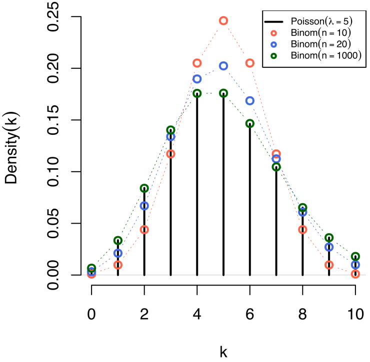
3. Let $X_i \sim P(\lambda_i)$ be n independent Poisson random variables. Then

$$X_1 + X_2 + \ldots + X_n \sim P(\lambda_1 + \lambda_2 + \ldots + \lambda_n).$$

3. Let $X_i \sim P(\lambda_i)$ be n independent Poisson random variables. Then
$$X_1 + X_2 + \ldots + X_n \sim P(\lambda_1 + \lambda_2 + \ldots + \lambda_n).$$

4. Suppose a Poisson random variable Z with mean μ represents the number of outcomes from some experiment. Let each outcome be independently classified in one of m categories, the kth of which occurs with probability p_k. Then the number of outcomes Z_k falling in category k is Poisson distributed with mean $\mu_k = p_k\mu$. Furthermore, the random variables $Z_1, \ldots, Z_m$ are independent.

3. If the number of trials n is large and the probability of a success p on a trial is small, then the total number of successes will be approximately a Poisson random variable with parameter $\lambda = np$. In the other word, $P(np) \approx B(n,p)$.



THE POISSON DISTRIBUTIONS $P(\lambda)$

EXAMPLE 7

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 mm. Find the probability of

A. exactly 2 flaws in 1 mm of wire.
B. exactly 10 flaws in 5 mm of wire.
C. at least 1 flaw in 0.5 mm of wire.

The Poisson distributions $P(\lambda)$

Example 7

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 mm. Find the probability of

- A. exactly 2 flaws in 1 mm of wire.
- B. exactly 10 flaws in 5 mm of wire.
- C. at least 1 flaw in 0.5 mm of wire.

Example 8

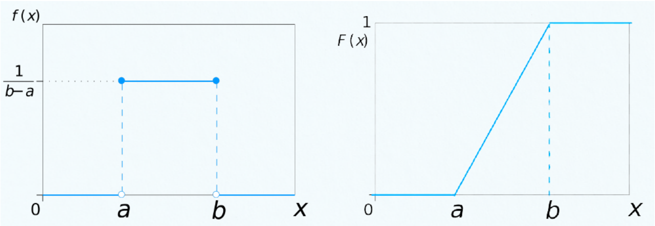
Suppose that 1 in 5000 light bulbs are defective. What is the probability that there are at least 3 defective light bulbs in a group of size 10000?

THE CONTINUOUS UNIFORM DISTRIBUTIONS $U(a,b)$

DEFINITION 2.1

A continuous uniform random variable has the pdf and cdf

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a,b] \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad F(x) = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & x \in [a,b) \\ 1, & x \geq b \end{cases}$$



THE CONTINUOUS UNIFORM DISTRIBUTIONS $U(a,b)$

PROPOSITION 2.1

If $X \sim U(a,b)$, then

1. $\mathbb{E}(X) = \dfrac{a+b}{2}$
2. $\mathbb{V}(X) = \dfrac{(b-a)^2}{12}$

The Continuous Uniform distributions $U(a,b)$

Proposition 2.1

If $X \sim U(a,b)$, then

1. $\mathbb{E}(X) = \dfrac{a+b}{2}$
2. $\mathbb{V}(X) = \dfrac{(b-a)^2}{12}$

Example 9

If X is uniformly distributed over (0, 10), calculate the probability that

A. $X < 3$
B. $X > 6$
C. $3 < X < 8$

The Continuous Uniform distributions $U(a,b)$

Example 10

A show is scheduled to start at 9:00 a.m., 9:30 a.m., and 10:00 a.m. Once the show starts, the gate will be closed. A visitor will arrive at the gate at a time uniformly distributed between 8:30 a.m. and 10:00 a.m. Find the probability that a visitor waits less than 10 minutes for a show.

THE NORMAL DISTRIBUTIONS $N(\mu,\sigma^2)$

DEFINITION 2.2

X is called to be of a normal distribution $N(\mu,\sigma^2)$ if its pdf or cdf satisfies

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}, x \in \mathbb{R},$$

PROPOSITION 2.2

If $X \sim N(\mu,\sigma^2)$, then

1. $\mathbb{E}(X) = \mu$, and $\mathbb{V}(X) = \sigma^2$.
2. The pdf of X is symmetric and has a bell-shaped curve centered at μ.

THE NORMAL DISTRIBUTIONS $N(\mu,\sigma^2)$

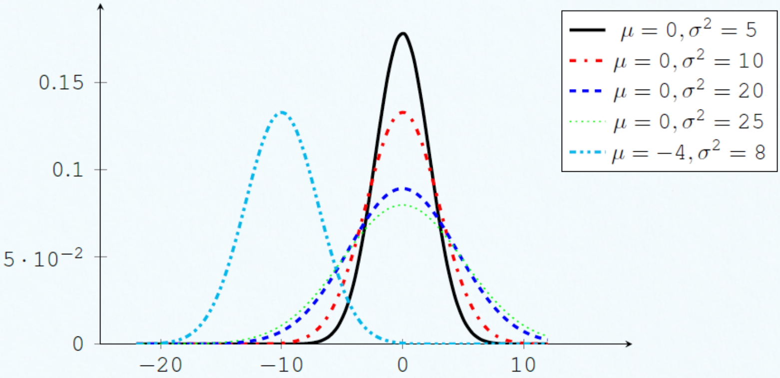


Figure: Pdf of $N(\mu,\sigma^2)$

The Normal distributions $N(\mu,\sigma^2)$

Definition 2.3

We often standardize a normal distribution $X \sim N(\mu,\sigma^2)$ by

$$Y = \frac{X-\mu}{\sigma} \sim N(0,1)$$

In this case, Y is called a random variable of standard normal distribution, or simply a standard score. Its pdf is

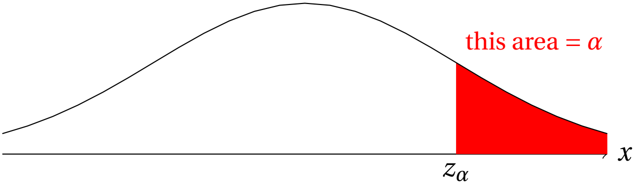
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, x \in \mathbb{R}$$

The Normal distributions $N(\mu,\sigma^2)$

The cdf of $X \sim N(0,1)$

$$\Phi(x) = \int_{-\infty}^{x} f(u)du$$

satisfies $\Phi(-x) = 1 - \Phi(x)$. The values of $\Phi(x)$ can be estimated by the Table of Standard Normal Distribution.
Denote z_α as the solution to $1 - \Phi(z) = \alpha$



z_α is called the upper α critical point or the $100(1-\alpha)$th percentile.

THE NORMAL DISTRIBUTIONS $N(\mu,\sigma^2)$

EXAMPLE 11

Suppose that the current measurements in a strip of wire follows a normal distribution with $\mu = 10$ and $\sigma = 2$ (mA).

A. What is the probability that the current measurement is between 9 and 11 mA?
B. Determine the value for which the probability that a current measurement is below this value is 0.98.

THE NORMAL DISTRIBUTIONS $N(\mu,\sigma^2)$

PROPOSITION 2.3

1. If $X \sim N(\mu,\sigma^2)$ and $Y = aX + b, a \neq 0$ then
$$Y \sim N(a\mu + b, a^2\sigma^2).$$
2. If $X_i \sim N(\mu_i,\sigma_i^2),\ i = 1,\dots,n$ and they are independent, then
$$\sum_{i=1}^{n} X_i \sim N\left(\sum_{i=1}^{n}\mu_i, \sum_{i=1}^{n}\sigma_i^2\right).$$

THE NORMAL DISTRIBUTIONS $N(\mu,\sigma^2)$

EXAMPLE 12

1. Let $X \sim N(5,9)$. What is the distribution of $Y = 2X - 6$?
2. Let $X_1 \sim N(2,5)$ and $X_2 \sim N(-3,4)$ be independent. Determine the distributions of the following random variables
   - A. $X = X_1 + X_2$.
   - B. $Y = X_1 - X_2$.
   - C. $Z = 3X_1 + 4X_2$.

The Exponential distributions $E(\lambda)$

Definition 2.4

A random variable T $(t > 0)$ has the Exponential distribution, denoted by, $T \sim E(\lambda)$, if its probability density function is

$$f(t) = \lambda e^{-\lambda t}, \quad t > 0, \tag{1}$$

where

- λ: the average number of events per unit length,
- t: is the number of unit-lengths until the next event occurs.

The Exponential distributions $E(\lambda)$

Definition 2.4

A random variable T $(t > 0)$ has the Exponential distribution, denoted by, $T \sim E(\lambda)$, if its probability density function is

$$f(t) = \lambda e^{-\lambda t}, \quad t > 0, \tag{1}$$

where

- λ: the average number of events per unit length,
- t: is the number of unit-lengths until the next event occurs.

Proposition 2.4

The random variable X that equals the distance between successive events from a Poisson process with mean number of events $\lambda > 0$ per unit length is an exponential random variable with parameter λ

The Exponential distributions $E(\lambda)$

Example 13

Suppose that there are 2 earthquakes each month in a certain country.

(A) Find the probability that there are at least 3 earthquakes in the next 2 month.

(B) Let T be the time (months) until the next earthquake occurs. Find $\mathbb{P}(T \leq 1)$.

THE EXPONENTIAL DISTRIBUTIONS $E(\lambda)$

PROPOSITION 2.5

If $T \sim E(\lambda)$ then

1. its cumulative distribution function is

$$F(t) = \mathbb{P}(T \leq t) = 1 - e^{-\lambda t}, \quad t > 0 \tag{2}$$

The Exponential distributions $E(\lambda)$

Proposition 2.5

If $T \sim E(\lambda)$ then

1. its cumulative distribution function is

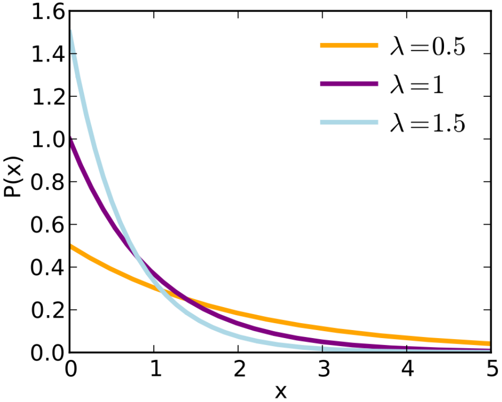
$$F(t) = \mathbb{P}(T \le t) = 1 - e^{-\lambda t}, \quad t > 0 \tag{2}$$

2.

$$\mathbb{E}(T) = \frac{1}{\lambda} \quad \text{and} \quad \mathbb{V}ar(T) = \frac{1}{\lambda^2}$$

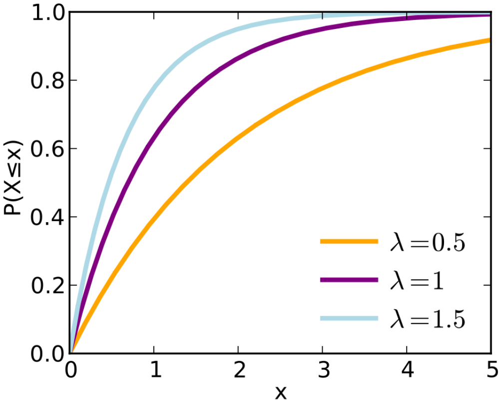
The Exponential distributions $E(\lambda)$

The p.d.f of $X \sim E(\lambda)$



The Exponential distributions $E(\lambda)$

The c.d.f of $X \sim E(\lambda)$



THE EXPONENTIAL DISTRIBUTIONS $E(\lambda)$

EXAMPLE 14

In a large corporate computer network, user log-ons to the system can be modeled as a Poisson process with a mean of 25 log-ons per hour.

1. What is the probability that 3 are no log-ons in an interval of 6 minutes?
2. What is the probability that the time until the next log-on is between 2 and 3 minutes?
3. Determine the interval of time such that the probability that no log-on occurs in the interval is 0.90.
4. Find the mean time until the next log-on.
5. Find the standard deviation of the time until the next log-on.

THE EXPONENTIAL DISTRIBUTIONS $E(\lambda)$

EXAMPLE 15

Let X denote the time between detections of a particle with a Geiger counter and assume that X has an exponential distribution with $E(X) = 1.4$ minutes.

(A) What is the probability that we detect a particle within 30 seconds of starting the counter.

(B) Suppose that we turn on the Geiger counter and wait three minutes without detecting a particle. What is the probability that a particle is detected in the next 30 seconds?

The central limit theorem

Theorem 2.1

Let $X_1, \ldots, X_n$ are n independent random variables and they have the same distribution with the mean μ and the finite variance σ^2.
Let $\overline{X} = \dfrac{X_1 + X_2 + \cdots + X_n}{n}$ denote the sample mean.

THE CENTRAL LIMIT THEOREM

THEOREM 2.1

Let $X_1,\ldots,X_n$ are n independent random variables and they have the same distribution with the mean μ and the finite variance σ^2. Let $\overline{X} = \dfrac{X_1+X_2+\cdots+X_n}{n}$ denote the sample mean. Then, the random variable

$$Z_n = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$$

converges in distribution to the standard normal r.v as n goes to infinity, that is

$$\lim_{n\to\infty} \mathbb{P}(Z_n \le x) = \Phi(x),\ \forall x \in \mathbb{R},$$

where $\Phi(x)$ is the standard normal CDF.

THE CENTRAL LIMIT THEOREM

THEOREM 2.1

Let $X_1,\ldots,X_n$ are n independent random variables and they have the same distribution with the mean μ and the finite variance σ^2.
Let $\overline{X} = \dfrac{X_1+X_2+\cdots+X_n}{n}$ denote the sample mean. Then, the random variable

$$Z_n = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$$

converges in distribution to the standard normal r.v as n goes to infinity, that is

$$\lim_{n\to\infty} \mathbb{P}(Z_n \le x) = \Phi(x),\ \forall x \in \mathbb{R},$$

where $\Phi(x)$ is the standard normal CDF.
Similarly, let $S_n = X_1+\ldots+X_n$, then $S_n \approx N\left(n\mu; n\sigma^2\right)$

Example 16 (Binomial distribution)

Assumptions:

- $X_1, X_2 \ldots$ are iid Bernoulli(p).
- $Z_n = \dfrac{X_1 + X_2 + \ldots + X_n - np}{\sqrt{np(1-p)}}$.

We choose $p = \frac{1}{3}$.

$$Z_1 = \frac{X_1 - p}{\sqrt{p(1-p)}}$$

PMF of Z_1

$$Z_2 = \frac{X_1 + X_2 - 2p}{\sqrt{2p(1-p)}}$$

PMF of Z_2

$$Z_3 = \frac{X_1 + X_2 + X_3 - 3p}{\sqrt{3p(1-p)}}$$

PMF of Z_3

$$Z_{30} = \frac{\sum_{i=1}^{30} X_i - 30p}{\sqrt{30p(1-p)}}$$

PMF of Z_{30}

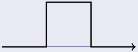
EXAMPLE 17

Assumptions:

- $X_1, X_2 \dots$ are iid Uniform(0,1).
- $Z_n = \dfrac{X_1 + X_2 + \dots + X_n - \frac{n}{2}}{\sqrt{\frac{n}{12}}}$.

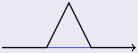
$$Z_1 = \frac{X_1 - \frac{1}{2}}{\sqrt{\frac{1}{12}}}$$

PDF of Z_1



$$Z_2 = \frac{X_1 + X_2 - 1}{\sqrt{\frac{2}{12}}}$$

PDF of Z_2



$$Z_3 = \frac{X_1 + X_2 + X_3 - \frac{3}{2}}{\sqrt{\frac{3}{12}}}$$

PDF of Z_3

$$Z_{30} = \frac{\sum_{i=1}^{30} X_i - \frac{30}{2}}{\sqrt{\frac{30}{12}}}$$

PDF of Z_{30}

NOTES

- The CLT does not matter what the distribution of the X_i's is. The X_i's can be discrete, continuous, or mixed random variables.

Notes

- The CLT does not matter what the distribution of the X_i's is. The X_i's can be discrete, continuous, or mixed random variables.
- The importance of the central limit theorem stems from the fact that, in many real applications, a certain random variable of interest is a sum of a large number of independent random variables.

THE CENTRAL LIMIT THEOREM

EXAMPLE 18

A bank teller serves customers standing in the queue one by one. Suppose that the service time X_i for customer i has mean $\mathbb{E}(X_i) = 2$ (minutes) and $\mathbb{V}(X_i) = 1$. We assume that service times for different bank customers are independent. Let Y be the total time the bank teller spends serving 50 customers. Find $P(90 < Y < 110)$.

EXAMPLE 19

A company producing cereals offers a toy in every 2 cereal package in celebration of their 50th anniversary. A father immediately buys 20 packages. Find the probability that he can get from 8 to 10 toys.

Continuity Correction for Discrete Random Variables

Let $X_1, X_2, \ldots, X_n$ be independent discrete random variables and let

$$Y = X_1 + X_2 + \cdots + X_n.$$

Suppose that we are interested in finding $\mathbb{P}(A) = \mathbb{P}(l \le Y \le u)$ using the CLT, where l and u are integers. Since Y is an integer-valued random variable, we can write

$$\mathbb{P}(A) = \mathbb{P}(l - 0.5 \le Y \le u + 0.5).$$

It turns out that the above expression sometimes provides a better approximation for P(A) when applying the CLT. This is called the continuity correction and it is particularly useful when X_i's are Bernoulli (i.e., $Y4$ is binomial).