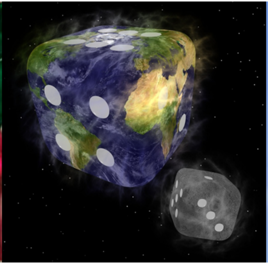
CHAPTER 5: CONTINUOUS RANDOM VARIABLES

Lecturer: Nguyen Minh Quan, PhD
Department of Mathematics
HCMC International University


CONTENTS

Introduction to continuous random variables

- In Chapter 4, we considered discrete random variables, that is, random variables whose set of possible values is either finite or countably infinite.
- There also exist random variables whose set of possible values is uncountable.
- Examples: The time that a train arrives at a specified stop and the lifetime of a transistor.

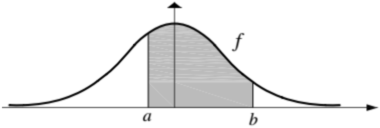
Introduction to continuous random variables

Definition

We say that X is a continuous random variable if there exists a nonnegative function f, defined for all real $x \in (-\infty, \infty)$, having the property that, for any set measurable B of real numbers,

$$P(X \in B) = \int_B f(x)\,dx$$

The function f is called the probability density function of the random variable X.



The shaded area under f is the probability that $X \in B = [a, b]$.

Properties of the probability density function

- $\int_{-\infty}^{\infty} f(x)\,dx = 1.$

  Because of this property, in general, if a function $g : \mathbb{R} \to [0, \infty)$ satisfies $\int_{-\infty}^{\infty} g(x)\,dx = 1$, we say that g is a probability density function, or simply a density function.

- $P(a \leqslant X \leqslant b) = \int\limits_a^b f(x)\,dx$

- $P(X = a) = 0.$

Continuous Random variables

Example

The amount of time, in hours, that a computer functions before breaking down is a continuous random variable with probability density function given by

$$f(x) = \begin{cases} \lambda e^{-x/100}, & x \geqslant 0 \\ 0, & x < 0 \end{cases}$$

What is the probability that

(a) a computer will function between 50 and 150 hours before breaking down,

(b) it will function less than 100 hours?

Continuous Random variables

Solution

(a)

$$1 = \int_{-\infty}^{\infty} f(x)dx = \lambda \int_{0}^{\infty} e^{-x/100} dx$$

$$1 = \lim_{t \to \infty} \left[-\lambda (100)\, e^{-x/100} \Big|_0^t \right] = 100\lambda \text{ or } \lambda = 1/100$$

$$P(50 < X < 150) = \int_{50}^{150} \frac{1}{100} e^{-x/100} dx = .384$$

(b)

$$P(X < 100) = \int_{0}^{100} \frac{1}{100} e^{-x/100} dx = .633$$

Continuous Random variables

Example

Suppose that X is a continuous random variable whose probability density function is given by

$$f(x) = \begin{cases} C\left(4x - 2x^2\right) & \text{if } 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) What is the value of C?
(b) Find $P(X > 1)$.

Relationship between F and f

Remark

The relationship between the cumulative distribution F and the probability density f is expressed by

$$F(x) = P(X \leqslant x) = \int_{-\infty}^{x} f(t)\,dt$$

- Differentiating both sides of the preceding equation yields

$$F'(x) = f(x)$$

- $f(x)$ is a measure of how likely it is that the random variable will be near x.

Continuous Random variables

Example

If X is continuous with distribution function F_X and density function f_X, find the density function of $Y = 2X$.

Solution

$$F_Y(x) = P(Y \leqslant x) = P(2X \leqslant x)$$
$$\rightarrow F_Y(x) = P(X \leqslant x/2) = F_X(x/2)$$

Differentiate to obtain

$$f_Y(x) = \frac{1}{2} f_X(x/2)$$

Expected value

If X is a continuous random variable having probability density function $f(x)$, then the expected value of X is defined by

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

Example

Find $E[X]$ when the density function of X is

$$f(x) = \begin{cases} 2x \text{ if } 0 \leqslant x \leqslant 1 \\ 0 \text{ otherwise} \end{cases}$$

$$E[X] = \int_{0}^{1} 2x^2 dx = 2/3$$

Expected value

Theorem

If X is a continuous random variable with probability density function $f(x)$, then for any real-valued function g,

$$E\left[g\left(X\right)\right] = \int_{-\infty}^{\infty} g\left(x\right) f\left(x\right) dx$$

Example

Find $E[e^X]$ when the density function of X is

$$f\left(x\right) = \begin{cases} 1 \text{ if } 0 \leqslant x \leqslant 1 \\ 0 \text{ otherwise} \end{cases}$$

$$E\left[e^X\right] = \int_0^1 e^x dx = e - 1.$$

Variance and standard deviation

- If a and b are constants, then $E[aX + b] = aE[X] + b$.
- How to measure the possible variation of X around $E(X) = \mu$?

$$Var(X) = E\left[(X - E(X))^2\right] = E\left(X^2\right) - \mu^2$$

$$Var(X) = \int_{-\infty}^{\infty} x^2 f(x)\, dx - \mu^2$$

- Standard deviation: $\sigma_X = \sqrt{Var(X)}$

Variance and standard deviation

Example

Find $Var(X)$ when the density function of X is

$$f(x) = \begin{cases} 2x \text{ if } 0 \leqslant x \leqslant 1 \\ 0 \text{ otherwise} \end{cases}$$

$$E\left(X^2\right) = \int_0^1 x^2 f(x)\, dx = \frac{1}{2}$$

$$\rightarrow Var(X) = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}$$

Variance and standard deviation

Properties

Let X be a random variable and k a real number. Then

$$\text{(i) } Var\left(X+k\right) = Var\left(X\right)$$

$$\text{(ii) } Var\left(kX\right) = k^2 Var\left(X\right)$$

Proof:

Examples

Compute Mean and Variance

A recent study has shown that airline passengers arrive at the gate with the amount of time (in hours) before the scheduled flight time given by the probability density function

$$f(t) = \frac{3}{4}\left(2t - t^2\right), \quad \text{for} \quad 0 \leqslant t \leqslant 2$$

a. Find and interpret the expected value for this distribution.
b. Compute the variance and the standard deviation.

Solution

a. The expected value is

$$E(T) = \mu = \int_0^2 t\left(\frac{3}{4}\right)\left(2t - t^2\right) dt = \frac{3}{4}\left(\frac{2t^3}{3} - \frac{t^4}{4}\right)\Bigg|_0^2 = 1.$$

Examples

Solution (Cont.)
This result indicates that passengers arrive at the gate an average of 1 hour before the scheduled flight time.

b. The variance is

$$Var(T) = \int_0^2 t^2 \left(\frac{3}{4}\right)\left(2t - t^2\right) dt - 1^2$$

$$Var(T) = \frac{3}{4}\left(\frac{t^4}{2} - \frac{t^5}{5}\right)\Bigg|_0^2 - 1 = \frac{1}{5}.$$

The standard deviation

$$\sigma = \frac{1}{\sqrt{5}}.$$

The Uniform Random Variables

Definition

A random variable is said to be uniformly distributed over the interval (α, β) if its probability density function is given by

$$f(x) = \begin{cases} \frac{1}{\beta-\alpha}, & \text{if } \alpha \leqslant x \leqslant \beta \\ 0, & \text{otherwise} \end{cases}$$

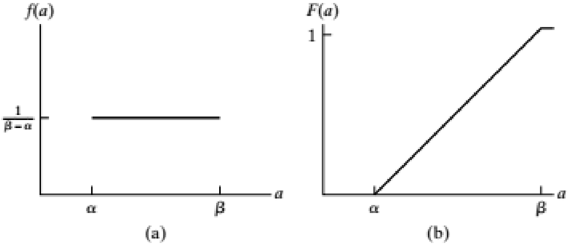


FIGURE 5.3: Graph of (a) $f(a)$ and (b) $F(a)$ for a uniform (α, β) random variable.

The Uniform Random Variables

Example

Calculate the cumulative distribution function of a random variable uniformly distributed over (α, β).

Solution:

We have $F(x) = \int\limits_{-\infty}^{x} f(x)\,dx$. This yields

$$F(x) = \begin{cases} 0, & x \leqslant \alpha \\ \dfrac{x-\alpha}{\beta-\alpha}, & \alpha < x < \beta \\ 1, & x \geqslant \beta \end{cases}$$

The Uniform Random Variables

Example

Let X be uniformly distributed over (α, β). Find $E(X)$ and $Var(X)$.

Solution:

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_{\alpha}^{\beta} \frac{x}{\beta - \alpha} dx = \frac{\alpha + \beta}{2}$$

$$Var(X) = E\left(X^2\right) - (E(X))^2$$

$$Var(X) = \int_{\alpha}^{\beta} \frac{x^2}{\beta - \alpha} dx - \left(\frac{\alpha + \beta}{2}\right)^2 = \frac{(\beta - \alpha)^2}{12}$$

The Uniform Random Variables

Example

If X is uniformly distributed over $(0, 10)$, calculate the probability that (a) $X < 3$, (b) $3 < X < 8$.

Solution:

$$P\left(X < 3\right) = \int_{0}^{3} \frac{1}{10} dx = \frac{3}{10}$$

$$P\left(3 < X < 8\right) = \int_{3}^{8} \frac{1}{10} dx = \frac{1}{2}$$

The Uniform Random Variables

Example

Buses arrive at a specified stop at 15-minute intervals starting at 7 A.M. That is, they arrive at 7, 7:15, 7:30, 7:45, and so on. If a passenger arrives at the stop at a time that is uniformly distributed between 7 and 7:30, find the probability that he waits
(a) less than 5 minutes for a bus;
(b) more than 10 minutes for a bus.

Solution Let X denote the number of minutes past 7 that the passenger arrives at the stop.
(a) Since X is a uniform random variable over the interval $(0, 30)$. The passenger will have to wait less than 5 minutes if (and only if) he arrives between 7:10 and 7:15 or between 7:25 and 7:30.

$$P(10 < X < 15) + P(25 < X < 30) = \frac{1}{3}.$$

(b) $P(0 < X < 5) + P(15 < X < 20) = 1/3$.

The Uniform Random Variables

Example

A stick of length 1 is split at a point U that is uniformly distributed over $(0, 1)$. Determine the expected length of the piece that contains the point p, $0 \leq p \leq 1$.

Solution:

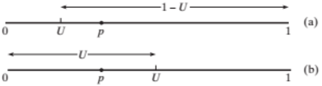


FIGURE 5.2: Substick containing point p: (a) $U < p$; (b) $U > p$.

Let $L_p(U)$ denote the length of the substick that contains the point p.

$$L_p(U) = \begin{cases} 1 - U & \text{if } U < p \\ U & \text{if } U > p \end{cases}$$

$$E[L_p(U)] = \int_0^1 L_p(U) f(u)\, du = \int_0^1 L_p(U)\, du = \frac{1}{2} + p(1-p).$$

The Uniform Random Variables

Example

The daily amount of coffee, in liters, dispensed by a machine located in an airport lobby is a random variable X having a continuous uniform distribution with $\alpha = 7$ and $\beta = 10$. Find the probability that on a given day the amount of coffee dispensed by this machine will be
(a) at most 8.8 liters;
(b) more than 7.4 liters but less than 9.5 liters;
(c) at least 8.5 liters

Hint: The probability density function is

$$f(x) = \frac{1}{\beta - \alpha} = \frac{1}{16}.$$

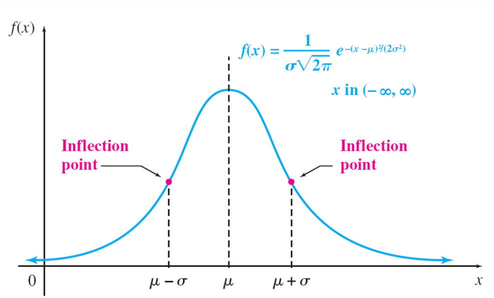
Normal Random Variables

Definition

X is a normal random variable (or simply that X is normally distributed, or X is a Gaussian random variable) with parameters μ and σ^2 if the density of X is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, -\infty < x < \infty$$

This density function is a bell-shaped curve that is symmetric around μ.



Normal Random Variables

Exercise

Show that

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2}dx = 1$$

Normal Random Variables

Exercise

Show that if X is normally distributed with parameters μ and σ^2 then $Y = \alpha X + \beta$ is normally distributed with parameters $\alpha\mu + \beta$ and $\alpha^2\sigma^2$.

Remark

If X is normally distributed with parameters μ and σ^2 then $Z = (X - \mu)/\sigma$ is normally distributed with parameters 0 and 1. Such a random variable Y is said to have the standard or unit normal distribution.

Normal Distributions

Preposition

Expected value and Variance are

$$E\left(X\right)=\mu, Var(X)=\sigma^2,$$

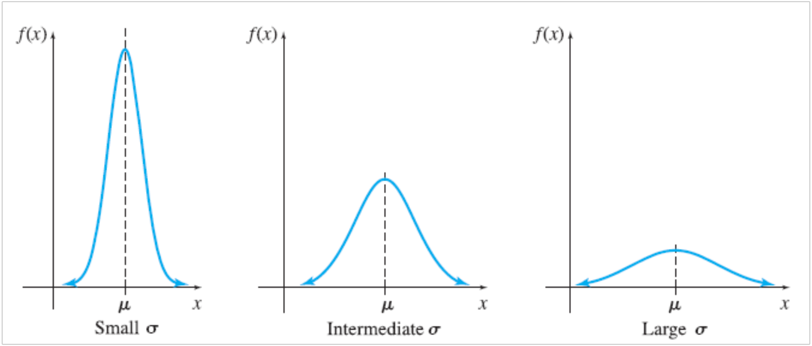
where is σ the standard deviation.

Probabilities for a normal distribution is given by

$$P(a\leqslant X\leqslant b)=\frac{1}{\sigma\sqrt{2\pi}}\int_a^b e^{-(x-\mu)^2/(2\sigma^2)}dx$$

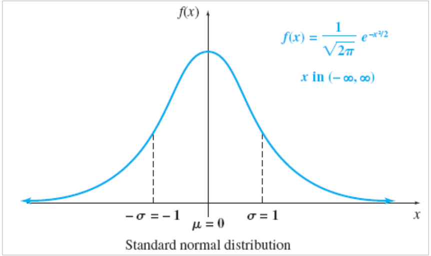
Normal Distributions

A larger value of σ produces a flatter normal curve, while smaller values of σ produce more values near the mean, resulting in a "taller" normal curve.



Standard Normal Distribution

The standard normal which has $\mu = 0$ and $\sigma = 1$ is called the standard normal distribution.



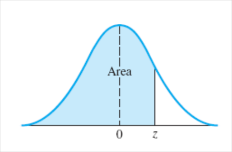
Note that

$$P(a \le X \le b) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \Phi(b) - \Phi(a)$$

where $\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$.

Standard Normal Distribution

$\Phi(z)$ is the area of the shaded region in the below figure:



The Table 5.1 (p.201) gives the values of $\Phi(z)$ for $z \geq 0$. Note that for $z \geq 0$, $\Phi(-z) = 1 - \Phi(z)$.

Example

If X is a standard normal random variable, then

$$P(0 \leq X \leq 1.5) = \Phi(1.5) - \Phi(0) = 0.9332 - 0.5 = 0.4332$$

$$P(-0.5 \leq X \leq 1.45) = \Phi(1.45) - \Phi(-0.5) = 0.9265 - 0.3085 = 0.6180$$

Method for other normal distributions but non-standard normal?

Standard Normal Distribution: Table 5.1 (p. 201)

TABLE 5.1: AREA $\Phi(x)$ UNDER THE STANDARD NORMAL CURVE TO THE LEFT OF X

X	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952

Standard Normal Distribution

z-Scores Theorem

Suppose a normal distribution has mean μ and standard deviation σ. The area under the associated normal curve that is to the left of the value x is exactly the same as the area to the left of

$$z = \frac{x - \mu}{\sigma}$$

for the standard normal curve.

Namely, if X is a normal random variable with mean μ and standard deviation σ then $Z = \frac{X-\mu}{\sigma}$ is a standard normal random variable.

Normal Distribution

Example

If X is a normal random variable with mean $\mu = 3$ and variance $\sigma^2 = 16$, find
(a) $P(X < 11)$.
(b) $P(X > -1)$.
(c) $P(2 < X < 7)$

Solution

$$(a) \qquad P(X < 11) = P\left(\frac{X-3}{4} < \frac{11-3}{4}\right) = P(Z < 2)$$

Therefore,

$$P(X < 11) = \Phi(2) = 0.9772$$

($Z = \dfrac{X-3}{4}$ is the standard normal random variable!)

Normal Distribution

Solution (Cont.)

$$(b) \qquad P\left(X > -1\right) = P\left(\frac{X-3}{4} > \frac{-1-3}{4}\right) = P\left(Z > -1\right)$$

Hence,

$$P\left(X > -1\right) = 1 - P\left(Z < -1\right) = 1 - 0.1587 = 0.8413$$

$$(c) \quad P\left(2 < X < 7\right) = P\left(\frac{2-3}{4} < \frac{X-3}{4} < \frac{7-3}{4}\right) = P\left(-1/4 < Z < 1\right)$$

$$P\left(2 < X < 7\right) = \Phi\left(1\right) - \Phi\left(-0.25\right) = 0.8413 - 0.4013 = 0.4400$$

(Note: $Z = \dfrac{X-3}{4}$ is standard normal!)

Normal Distribution

Example

An expert witness in a paternity suit testifies that the length (in days) of human gestation is approximately normally distributed with parameters $\mu = 270$ and $\sigma^2 = 100$. The defendant in the suit is able to prove that he was out of the country during a period that began 290 days before the birth of the child and ended 240 days before the birth. If the defendant was, in fact, the father of the child, what is the probability that the mother could have had the very long or very short gestation indicated by the testimony?

Solution Let X denote the length of the gestation, and assume that the defendant is the father. Then the probability that the birth could occur within the indicated period is

$$P(X > 290 \text{ or } X < 240) = P(X > 290) + P(X < 240)$$

$$= P\left(\frac{X-270}{10} > \frac{290-270}{10}\right) + P\left(\frac{X-270}{10} < \frac{240-270}{10}\right) \approx 0.0241$$

Normal Distribution

Example: Life Spans

According to actuarial tables, life spans in the United States are approximately normally distributed with a mean of about 75 years and a standard deviation of about 16 years. By computing the areas under the associated normal curve, find the following probabilities.
(a) Find the probability that a randomly selected person lives less than 88 years.
(b) Find the probability that a randomly selected person lives more than 67 years.

Solution (a) Let T represent the life span of a random individual. We need to find

$$P(T < 88) = P\left(\frac{T-\mu}{\sigma} < \frac{88-\mu}{\sigma}\right) = P\left(\frac{T-75}{16} < \frac{88-75}{16}\right)$$

$$\text{Thus} \qquad P(T < 88) = P(Z < 0.81) = \Phi(0.81) = 0.7910$$

(Z is standard normal!)

Normal Distribution

Solution (Cont.)

(b) The probability that a randomly selected person lives more than 67 years is $P(T > 67)$.

$$P(T > 67) = P\left(\frac{T-\mu}{\sigma} > \frac{67-\mu}{\sigma}\right) = P\left(\frac{T-75}{16} > \frac{67-75}{16}\right)$$

$$P(T > 67) = P(Z > -0.5) = 1 - P(Z < -0.5) = 1 - 0.3085 = 0.6915$$

(Z is standard normal!)

Normal Distribution

Example

Suppose that a Scottish soldier's chest size is normally distributed with mean 39.8 and standard deviation 2.05 inches, respectively. What is the probability that of 20 randomly selected Scottish soldiers, five have a chest of at least 40 inches?
(a) Find the probability that a randomly selected Scottish soldier has a chest of 40 or more inches.
(b) What is the probability that of 20 randomly selected Scottish soldiers, five have a chest of at least 40 inches?.

Hint

$$(a)\ \mathrm{P}\left(\mathrm{X} \geq 40\right) = P\left(Z \geq \frac{40-39.8}{2.05}\right) = P\left(Z \geq 0.1\right) = 1 - \Phi\left(0.1\right) = 0.46$$

$$(b)\quad \binom{20}{5} p^5 \left(1-p\right)^{15} = \binom{20}{5} 0.46^5 \left(1-0.46\right)^{15} \approx 0.03$$

Normal Distribution

Exercise

The average annual amount American households spend for daily transportation is \$6312 (Money, August 2001). Assume that the amount spent is normally distributed.

(a) Suppose you learn that 5% of American households spend less than \$1000 for daily transportation. What is the standard deviation of the amount spent?

(b) What is the probability that a household spends between \$4000 and \$6000?

(c) What is the range of spending for the 3% of households with the highest daily transportation cost?

Normal Distribution

Exercise

A soft-drink machine is regulated so that it discharges an average of 200 milliliters per cup. If the amount of drink is normally distributed with a standard deviation equal to 15 milliliters,

(a) what fraction of the cups will contain more than 224 milliliters?

(b) what is the probability that a cup contains between 191 and 209 milliliters?

(c) how many cups will probably overflow if 230 milliliter cups are used for the next 1000 drinks?

(d) below what value do we get the smallest 25% of the drinks?

The Normal Approximation to the Binomial Distribution

The DeMoivre-Laplace limit theorem

If S_n denotes the number of successes that occur when n independent trials, each resulting in a success with probability p, are performed, then, for any $a < b$

$$P\left(a \leqslant \frac{S_n - np}{\sqrt{np\left(1-p\right)}} \leqslant b\right) \rightarrow \Phi\left(b\right) - \Phi\left(a\right)$$

as $n \rightarrow \infty$.

The Normal Approximation to the Binomial Distribution

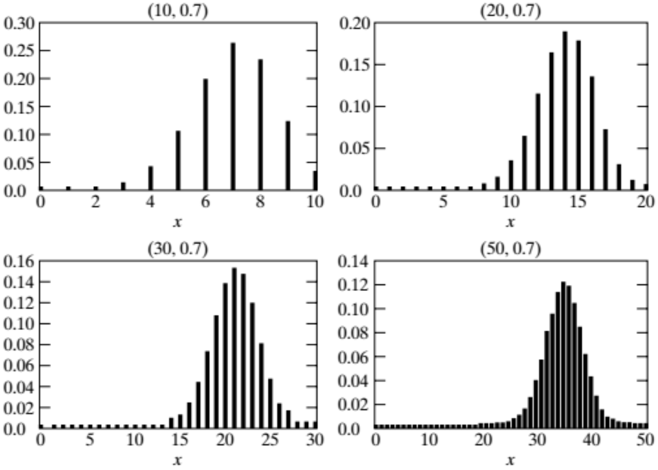


FIGURE 5.6: The probability mass function of a binomial (n, p) random variable becomes more and more "normal" as n becomes larger and larger.

The Normal Approximation to the Binomial Distribution

Example

Let X be the number of times that a fair coin that is flipped 40 times lands on heads. Find the probability that $X = 20$. Use the normal approximation and then compare it with the exact solution.

Solution

$$P(X = 20) = P(19.5 \leqslant X \leqslant 20.5)$$
$$= P\left(\frac{19.5 - 20}{\sqrt{10}} < \frac{X - 20}{\sqrt{10}} < \frac{20.5 - 20}{\sqrt{10}}\right)$$
$$= \Phi(0.16) - \Phi(-0.16) = 0.1272$$
$$P(X = 20) = \begin{pmatrix} 40 \\ 20 \end{pmatrix} \left(\frac{1}{2}\right)^{40} \approx 0.1254$$

The Normal Approximation to the Binomial Distribution

Example

The ideal size of a first-year class at a particular college is 150 students. The college, knowing from past experience that, on the average, only 30 percent of those accepted for admission will actually attend, uses a policy of approving the applications of 450 students. Compute the probability that more than 150 first-year students attend this college.

Solution

Let X denote the number of students that attend; then assuming that each accepted applicant will independently attend, it follows that X is a binomial random variable with parameters $n = 450$ and $p = 0.3$.

$$P\left(X > 150.5\right) = P\left(\frac{X - (450)(0.3)}{\sqrt{450\,(0.3)\,(0.7)}} \geqslant \frac{150.5 - (450)(0.3)}{\sqrt{450\,(0.3)\,(0.7)}}\right)$$

$$\approx P\left(Z > 1.59\right) = 0.06 \text{ (Z is standard normal)}$$

The Normal Approximation to the Binomial Distribution

Example

When you sign up for a credit card, do you read the contract carefully? In a FindLaw.com survey, individuals were asked, "How closely do you read a contract for a credit card?" (USA Today, 2003). The findings were that 44% read every word, 33% read enough to understand the contract, 11% just glance at it, and 4% don't read it at all.

(a) For a sample of 500 people, how many would you expect to say that they read every word of a credit card contract?

(b) For a sample of 500 people, what is the probability that 200 or fewer will say they read every word of a credit card contract?

(c) For a sample of 500 people, what is the probability that at least 15 say they don't read credit card contracts?

Exponential distribution

Definition

A continuous random variable whose probability density function is given, for some $\lambda > 0$

$$f(x) = \begin{cases} \lambda e^{-\lambda x} \text{ if } x \geqslant 0 \\ 0 \text{ if } x < 0 \end{cases}$$

is said to be an exponential random variable (or, more simply, is said to be exponentially distributed) with parameter λ.

The cumulative distribution function $F(x)$ of an exponential random variable is given by

$$F(x) = 1 - e^{-\lambda x}, \quad x \geqslant 0.$$

Exponential distribution

Theorem

If X be an exponential random variable, then

$$F(x) = 1 - e^{-\lambda x}, x \geqslant 0$$

$$E[X] = \frac{1}{\lambda}, \quad Var(X) = \frac{1}{\lambda^2}$$

Proof instructions:

Calculate the moment generating function:

$$\phi(t) = E\left[e^{tX}\right] = \frac{\lambda}{\lambda - t}, \quad \text{for } t < \lambda$$

Exponential distribution

Example

Suppose the useful life (in hours) of a hearing aid battery is the random variable t, with probability density function given by the exponential distribution

$$f(t) = \frac{1}{20} e^{-t/20}, \qquad t \geq 0$$

Find the probability that a particular battery, selected at random, has a useful life of less than 100 hours.

Solution The probability is given by

$$P(T \leq 100) = \int_0^{100} \frac{1}{20} e^{-t/20} dt = \frac{1}{20} \left(-20e^{-t/20} \right) \Big|_0^{100} = 0.9933$$

Exponential distribution

Example

Suppose that every three months, on average, an earthquake occurs in California. What is the probability that the next earthquake occurs after three but before seven months?

Solution:

Let X be the time (in months) until the next earthquake; it can be assumed that X is an exponential random variable with $\lambda = 1/3$.

$$P(3 < X < 7) = F(7) - F(3) = \left(1 - e^{-7/3}\right) - \left(1 - e^{-1}\right) \simeq 0.27$$

Exponential distribution

Exercise

At an intersection there are two accidents per day, on average. What is the probability that after the next accident there will be no accidents at all for the next two days?

Solution: Let X be the time (in days) between the next two accidents. It can be assumed that X is exponential with parameter $\lambda = 2$.

$$P(X > 2) = 1 - F(2) = e^{-4}.$$

Exponential distribution

Exercise

At an intersection there are two accidents per day, on average. What is the probability that after the next accident there will be no accidents at all for the next two days?

Hint: Let X be the time (in days) between the next two accidents. It can be assumed that X is exponential with parameter $\lambda = 2$.

Exponential distribution: Memoryless property

Memoryless property

The key property of an exponential random variable is that it is memoryless, where we say that a nonnegative random variable X is memoryless if

$$P\left(X > s + t \middle| X > t\right) = P\left(X > s\right), \quad \forall s, t \geqslant 0$$

Theorem

Exponentially distributed random variables are memoryless.

Proof instructions: Note that $P\left(X > x\right) = e^{\lambda x}, x > 0$. Thus, $P\left(X > s + t\right) = P\left(X > t\right) P\left(X > s\right)$.

Exponential distribution: Memoryless property

Example

Suppose that a number of miles that a car can run before its battery wears out is exponentially distributed with an average value of 10,000 miles.
(a) If a person desires to take a 5,000-mile trip, what is the probability that she will be able to complete her trip without having to replace her car battery?
(b) What can be said when the distribution is not exponential?

Solution:
(a) We note that the remaining lifetime (in thousands of miles) of the battery is exponential with parameter $\lambda = 1/10$.

$$P\left(remaining\ \ lifetime > 5\right) = 1 - F\left(5\right) = e^{-5\lambda} = e^{-1/2} \approx 0.604.$$

Exponential distribution: Memoryless property

(b) If the lifetime distribution F is not exponential, then the relevant probability is

$$P\left(\mathit{lifetime} > t+5 \middle| \mathit{lifetime} > t \right) = \frac{1 - F(t+5)}{F(t)}$$

where t is the number of miles used.

Exercise

The lifetime of a TV tube (in years) is an exponential random variable with mean 10. If Jim bought his TV set 10 years ago, what is the probability that its tube will last another 10 years?

-END OF CHAPTER 5-