Basic Statistical Terms

Statistics: The science of using information discovered from collecting, organizing, and studying numbers(data).

Statistics is the discipline that concerns the collection, organization, displaying, analysis, interpretation, and presentation of data.

Data: Numbers, letters, or special characters representing measurements of the properties of one's analytic units, or cases, in a study; data are the raw material of statistics.

Descriptive Statistics: They are brief descriptive coefficients that summarize a given data set.

Inferential Statistics: The body of statistical techniques concerned with making inferences about a population based on drawing a sample from it.

Population: The collection of all the elements of interest.

Sample is the smaller part of the whole i.e(= that is, yani anlamında) a subset of the entire population.

Sample Space: The set of all outcomes of an experiment.

Parameter: A summary measure of some characteristic for the population, such as the population mean or proportion.

A statistic is defined as a numerical value, which is obtained from a sample of data.

The distribution of a statistical data set (or a population) is a listing or function showing all the possible values (or intervals) of the data and how often they occur.

The mean (average) of a data set is found by adding all numbers in the data set and then dividing by the number of values in the set.

The median is the middle value when a data set is ordered from least to greatest.

The mode is the number that occurs most often in a data set.

Variance: The avarage of the squarred differences from the mean.

Standard Deviation: The square root of a variable's variance. The standard deviation is the most commonly used measure of dispersion and represents approximately the average distance of values from the mean of a distribution.

Week 3- 18/03/2021

Topic: Continutiy of Basic Statistical Terms

Variable

Generally, it is any quantity, that varies. The characteristic measured or observed when an experiment is carried out or an observation is made. Variables may be non-numerical or numerical. Since a non-numerical observation can always be coded numerically, a variable is usually taken to be numerical.

Category

It is a homogeneous class or group of a population of objects or measurements.

Categorical Variable / Qualitative Variable.

A variable that denotes quality rather than a quantity that can be measured on a scale.

Quantitative Variable / Numerical Variable

A variable that takes numerical values for which arithmetic makes sense, for example, counts, temperatures, weights, amounts of money, *etc.* For some variables that take numerical values, arithmetic with those values does not make sense; such variables are not quantitative. For example, adding and subtracting social security numbers does not make sense. Quantitative variables typically have units of measurement, such as inches, people, or pounds

Random Variable

A random variable is an assignment of numbers to possible outcomes of a random experiment. For example, consider tossing three coins. The number of heads showing when the coins land is a random variable: it assigns the number 0 to the outcome {T, T, T}, the number 1 to the outcome {T, T, H}, the number 2 to the outcome {T, H, H}, and the number 3 to the outcome {H, H, H}.

Probabilities are the study of *"chance"*. When we *calculate the probability* of something occurring we are *calculating the likelihood of it happening*.

Observation

The act of watching somebody or somothing carefully, especially to learn something.

Experiment

A process by which an observation or outcome is obtained.

In *probabilities*, an ***experiment*** is a process (could be "anything") in which there are one or more (usually more) possible ***outcomes*** each of which depends on *chance*.

Trial

Single performance of well-defined experiment.

Outcome

An outcome is the result of an experiment or other situation involving uncertainity.

Sample Space

The set S of all possible outcomes of an experiment is called Sample Space.

The *sample space* is usually written, or illustrated, using one of the following:

- a list of *all the possible outcomes* written inside a *set* that we call S,
- a *sample space diagram*, or
- a *Venn Diagram*.

Event

Any subset E of the sample space S.

Given an *experiment*, along with its possible *outcomes*, an *event* is the name given to either one of the possible *outcomes*, or a *group of outcomes*.

Events are usually referred to using a capital letter, such as A, B, C, ...

Union of Events

The occurance of either of two(or more) events.

Intersection of Events

The joint occurance of two or more events.

Exculusive Events

Exculusive events are the events whose interaction or the sample space that these events occuring at the same time, is empty set.

Probability

A probability is a number expressed as either:

- a Decimal
- a Fraction
- a Percentage

<u>It's value is a measure of the likelihood of an event occurring.</u>

<u>A quantitative measure of uncertainty.</u>

Notation

Given an event A, the probability of event A occurring is written:

$$p(A)$$

Read: "the probability of event A"

The likelihood of an event A occuring is measured on a scale that goes from 0 to 1, where:

- 0 is the probability of somthing impossible.
- 1 is the probability of something certain.

All other events have a probability that lies somewhere in between these two values.

Axioms of Probability. / Kolmogorov Axioms

There are three axioms of probability:

(1) Chances are always at least zero.

(2) The chance that *something* happens is 100%.

(3) If two events cannot both occur at the same time, the chance that either one occurs is the sum of the chances that each occurs.

For example, consider an experiment that consists of tossing a coin once.

- The first axiom says that the chance that the coin lands heads, for instance, must be at least zero.
- The second axiom says that the chance that the coin either lands heads or lands tails or lands on its edge or doesn't land at all is 100%.
- The third axiom says that the chance that the coin either lands heads or lands tails is the sum of the chance that the coin lands heads and the chance that the coin lands tails, because both cannot occur in the same coin toss.

All other mathematical facts about probability can be derived from these three axioms. For example, it is true that the chance that an event does not occur is (100% − the chance that the event occurs). This is a consequence of the second and third axioms.

Definition

(Kolmogorov Axioms) Consider a random experiment with sample space S and an event $A \subset S$ of interest. If $P(A)$ is defined and if

Axiom 1. $P(A) \geq 0$

Axiom 2. $P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$ *where* $A_1, A_2, \ldots$ *are disjoint events, and*

Axiom 3. $P(S) = 1$

then $P(A)$ is the probability of event A occurring.

Joint Probability

The probabity of the intersection of events.

Conditional Probability

The probability that an event occurs when the outcome of some other event is given. In probability theory, conditional probability is a measure of the probability of an event occurring, given that another event (by assumption, presumption, assertion or evidence) has already occurred.

Notation: P(A\B).

Independence

In the calculus of probabilities, independence is usually defined by reference to the principle of compound probabilites. Two events are independent if the probability of the one is the same whether the other is given or not.

P(A) = P(A\B) and P(B) =P(B\A)

Mesleki Yabancı Dil: 25 Mart 2021 (4.Hafta)

Axiom (Tr. Aksiyom)

An axiom, postulate or assumption is a statement that is taken to be true, to serve as a premise or starting point for further reasoning and arguments.

Complementary event (Tr. Tümleyen Olay)

The complementary event A' to an event A is the event 'A does not occur'.

With each event A is associated the complementary event A' consisting of those experimental outcomes that do not belong to A.

Continuos Random Variable

A variable whose set of possible values is a continuous interval of real numbers x, such that a<x<b, in which a can be $-\infty$ ∞ and b can be ∞ .

A probability **distribution is sometimes said to be continuous** when it relates to a continuous random variable.

In probability theory, a **probability density function (PDF**), or density of a continuous random variable, is a function whose value at any given sample (or point) in the sample space (the set of possible values taken by the random variable) can be interpreted as providing a relative likelihood that the value of the random variable would equal that sample.

For a continuous random variable X the probability density function f is such that

$P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x)dx$ for all x1<x2.

Discrete Random Variable

A random variable whose set of possible values is a finite or infinite sequence of numbers x1, x2,….

The probability distribution of a discrete random variable is referred to as a **discrete distribution.**

For a discrete random variable X, with possible values x1, x2,…, the function f, defined by

$f(x_j) = P(X = x_j), j = 1,2, \ldots$ is the **probability function of X.**

Cumulative Distribution Function

In probability theory and statistics, the cumulative distribution function (CDF) of a real-valued random variable X, or just distribution function of X, evaluated at x, is the probability that X will take a value less than or equal to x.

$$F_x(x) = P(X \leq x)$$

Expected Value

The expected value of a random variable is the long-term limiting average of its values in independent repeated experiments. The expected value of the random variable X is denoted by E(X).

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

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

Variance

The variance of a random variable X, Var(X), is the expected value of the squared difference between the variable and its expected value: Var(X) = E((X − E(X))^2)= E(x^2)-(E(x))^2

Standart Deviation

In statistics, the standard deviation is a measure of the amount of variation or dispersion of a set of values.A low standard deviation indicates that the values tend to be close to the mean (also called the expected value) of the set, while a high standard deviation indicates that the values are spread out over a wider range.

Most commonly represented in mathematical texts and equations by the lower case Greek letter sigma σ, for the population standard deviation, or the Latin letter s, for the sample standard deviation.

Standard Error of a Statistic

The standard error (SE)of a statistic (usually an estimate of a parameter) is the standard deviation of its sampling distribution or an estimate of that standard deviation. If the statistic is the sample mean, it is called the standard error of the mean (SEM).

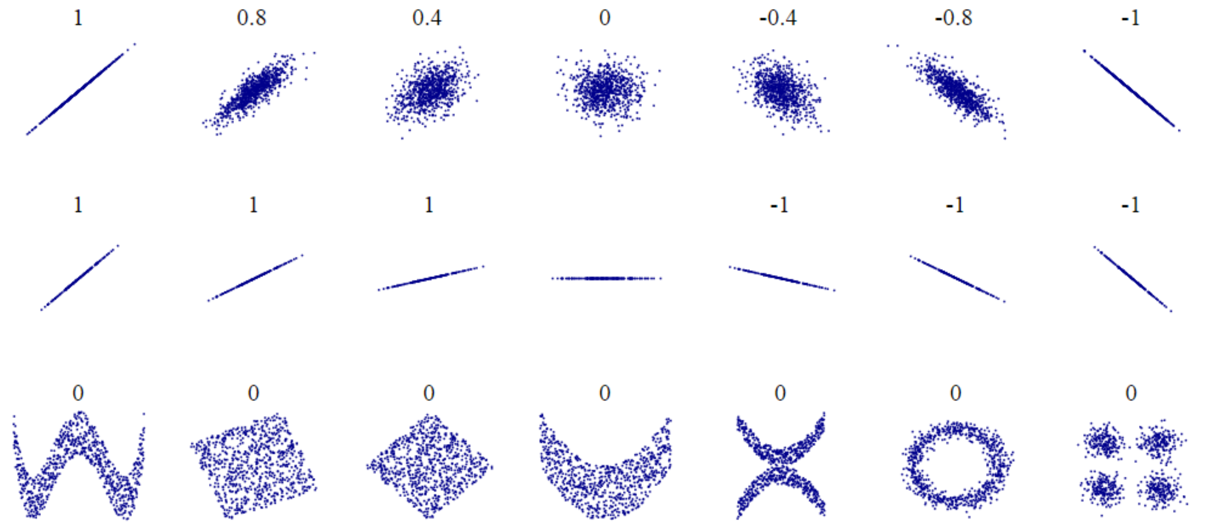
Covariance

In probability theory and statistics, covariance is a measure of the joint variability of two random variables. If the greater values of one variable mainly correspond with the greater values of the other variable, and the same holds for the lesser values (that is, the variables tend to show similar behavior), the covariance is positive.In the opposite case, when the greater values of one variable mainly correspond to the lesser values of the other, (that is, the variables tend to show opposite behavior), the covariance is negative.

$$\operatorname{cov}(X, Y) = \operatorname{E}\left[(X - \operatorname{E}[X])(Y - \operatorname{E}[Y])\right]$$

Correlation

In statistics, correlation or dependence is any statistical relationship, whether causal or not, between two random variables or bivariate data. In the broadest sense correlation is any statistical association, though it commonly refers to the degree to which a pair of variables are linearly related. Familiar examples of dependent phenomena include the correlation between the height of parents and their offspring, and the correlation between the price of a good and the quantity the consumers are willing to purchase, as it is depicted in the so-called demand curve.



$$\rho_{X,Y} = \mathrm{corr}(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\mathrm{E}[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}$$

Moments of a Random Variable

The "moments" of a random variable (or of its distribution) are expected values of powers or related functions of the random variable.

In probability theory and statistics, a central moment is a moment of a probability distribution of a random variable about the random variable's mean; that is, it is the expected value of a specified integer power of the deviation of the random variable from the mean.

The nth moment about the mean (or nth central moment) of a real-valued random variable X is the quantity μ_n := E[(X − E[X])^n], where E is the expectation operator.

The nth moment about zero : E[(X-0)^n].

Skewness - Çarpıklık

In probability theory and statistics, skewness is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean. The skewness value can be positive, zero, negative, or undefined.

The third central moment is the measure of the lopsidedness of the distribution; any symmetric distribution will have a third central moment, if defined, of zero. The normalised third central moment is called the skewness, often γ.

Kurtosis – Basıklık

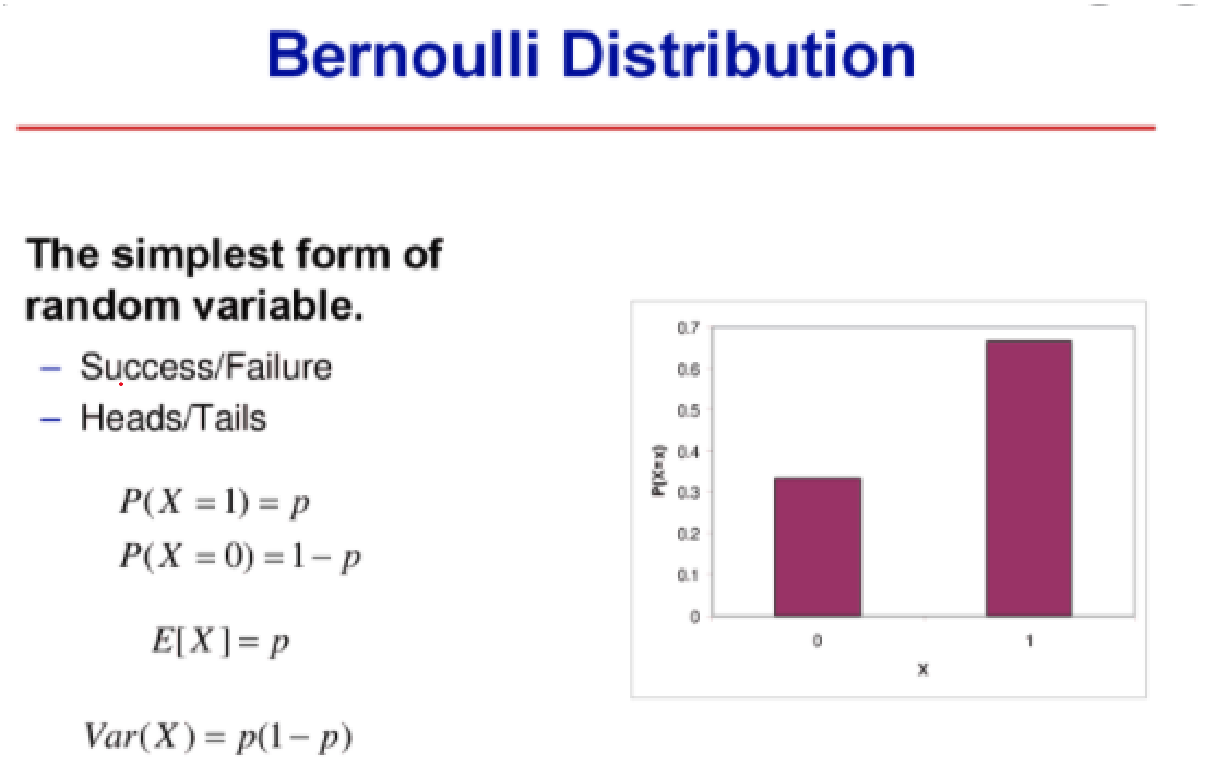
In probability theory and statistics, kurtosis (from Greek: κυρτός, kyrtos or kurtos, meaning "curved, arching") is a measure of the "tailedness" of the probability distribution of a real-valued random variable. Like skewness, kurtosis describes the shape of a probability distribution and there are different ways of quantifying it for a theoretical distribution and corresponding ways of estimating it from a sample from a population. Different measures of kurtosis may have different interpretations.

MESLEKİ YABANCI DİL: 1 NİSAN 2021 (5.HAFTA)

The probability distribution of a discrete random variable is referred to as a **discrete distribution.**

THE BERNOULLİ DİSTRİBUTİON

In probability theory and statistics, **the Bernoulli distribution**, named after Swiss mathematician Jacob Bernoulli, is the discrete probability distribution of a random variable which takes the value 1 with probability p and the value 0 with probability q=1-p. Less formally, it can be thought of as a model for the set of possible outcomes of any single experiment that asks a yes–no question. Such questions lead to outcomes that are boolean-valued: a single bit whose value is success/yes/true/one with probability p and failure/no/false/zero with probability q. It can be used to represent a (possibly biased) coin toss where 1 and 0 would represent "heads" and "tails" (or vice versa), respectively, and p would be the probability of the coin landing on heads or tails, respectively.



THE BİNOMİAL DİSTRİBUTİON

In probability theory and statistics, **the binomial distribution** with parameters n and p is the discrete probability distribution of the number of successes in a sequence of n independent experiments, each asking a yes–no question, and each with its own Boolean-valued outcome: success (with probability p) or failure (with probability q = 1 − p). A single success/failure experiment is also called a Bernoulli trial or Bernoulli experiment, and a sequence of outcomes is called a Bernoulli process; for a single trial, i.e., n = 1, the binomial distribution is a Bernoulli distribution.

Binomial Distribution Formula

$$P(x) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{(n-x)!\,x!} p^x q^{n-x}$$

where

n = the number of trials (or the number being sampled)

x = the number of successes desired

p = probability of getting a success in one trial

q = 1 − *p* = the probability of getting a failure in one trial

Historical Note

Independent trials having a common probability of success p were first studied by the Swiss mathematician Jacques Bernoulli (1654–1705). In his book Ars Conjectandi (The Art of Conjecturing), published by his nephew Nicholas eight years after his death in 1713, Bernoulli showed that if the number of such trials were large, then the proportion of them that were successes would be close to p with a probability near 1. Jacques Bernoulli was from the first generation of the most famous mathematical family of all time. Altogether, there were between 8 and 12 Bernoullis, spread over three generations, who made fundamental contributions to probability, statistics, and mathematics. One difficulty in knowing their exact number is the fact that several had the same name. (For example, two of the sons of Jacques's brother Jean were named Jacques and Jean.) Another difficulty is that several of the Bernoullis were known by different names in different places. Our Jacques (sometimes written Jaques) was, for instance, also known as Jakob (sometimes written Jacob) and as James Bernoulli. But whatever their number, their influence and output were prodigious. Like the Bachs of music, the Bernoullis of mathematics were a family for the ages!

THE POISSON RANDOM VARIABLE / DİSTRİBUTİON

A random variable X that takes on one of the values 0, 1, 2, ... is said to be a Poisson random variable with parameter λ if, for some λ > 0,

$$p(i) = P\{X = i\} = e^{-\lambda}\frac{\lambda^i}{i!} \quad i = 0, 1, 2, \ldots \qquad (7.1)$$

Equation (7.1) defines a probability mass function, since

$$\sum_{i=0}^{\infty} p(i) = e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{-\lambda}e^{\lambda} = 1$$

The Poisson probability distribution was introduced by Simeon Denis Poisson in a ´ book he wrote regarding the application of probability theory to lawsuits, criminal trials, and the like. This book, published in 1837, was entitled Recherches sur la probabilite des jugements en mati ´ ere criminelle et en mati ` ere civile (Investigations into the Probability of Verdicts in Criminal and Civil Matters). The Poisson random variable has a tremendous range of applications in diverse areas because it may be used as an approximation for a binomial random variable with parameters (n, p) when n is large and p is small enough so that np is of moderate size. To see this, suppose that X is a binomial random variable with parameters (n, p), and let λ = np. Then

$$\begin{aligned} P\{X = i\} &= \frac{n!}{(n-i)!i!} p^i (1-p)^{n-i} \\ &= \frac{n!}{(n-i)!i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i} \\ &= \frac{n(n-1)\cdots(n-i+1)}{n^i} \frac{\lambda^i}{i!} \frac{(1-\lambda/n)^n}{(1-\lambda/n)^i} \end{aligned}$$

Now, for n large and λ moderate,

$$\left(1 - \frac{\lambda}{n}\right)^n \approx e^{-\lambda} \quad \frac{n(n-1)\cdots(n-i+1)}{n^i} \approx 1 \quad \left(1 - \frac{\lambda}{n}\right)^i \approx 1$$

Hence, for n large and λ moderate,

$$P\{X = i\} \approx e^{-\lambda}\frac{\lambda^i}{i!}$$

In other words, if n independent trials, each of which results in a success with probability p, are performed, then, when n is large and p is small enough to make np moderate, the number of successes occurring is approximately a Poisson random variable with parameter λ = np. This value λ (which will later be shown to equal the expected number of successes) will usually be determined empirically. Some examples of random variables that generally obey the Poisson probability law [that is, they obey Equation (7.1)] are as follows:

1. The number of misprints on a page (or a group of pages) of a book
2. The number of people in a community who survive to age 100
3. The number of wrong telephone numbers that are dialed in a day
4. The number of packages of dog biscuits sold in a particular store each day
5. The number of customers entering a post office on a given day
6. The number of vacancies occurring during a year in the federal judicial system
7. The number of α-particles discharged in a fixed period of time from some radioactive material

THE GEOMETRIC RANDOM VARIABLE / DİSTRİBUTİON

4.8.1 The Geometric Random Variable

Suppose that independent trials, each having a probability $p, 0 < p < 1$, of being a success, are performed until a success occurs. If we let X equal the number of trials required, then

$$P\{X = n\} = (1 - p)^{n-1}p \qquad n = 1, 2, \ldots \tag{8.1}$$

Equation (8.1) follows because, in order for X to equal n, it is necessary and sufficient that the first $n - 1$ trials are failures and the nth trial is a success. Equation (8.1) then follows, since the outcomes of the successive trials are assumed to be independent.

Since

$$\sum_{n=1}^{\infty} P\{X = n\} = p\sum_{n=1}^{\infty}(1 - p)^{n-1} = \frac{p}{1 - (1 - p)} = 1$$

it follows that, with probability 1, a success will eventually occur. Any random variable X whose probability mass function is given by Equation (8.1) is said to be a *geometric* random variable with parameter p.

EXAMPLE 8b

Find the expected value of a geometric random variable.

Solution. With $q = 1 - p$, we have

$$\begin{aligned}
E[X] &= \sum_{i=1}^{\infty} iq^{i-1}p \\
&= \sum_{i=1}^{\infty}(i - 1 + 1)q^{i-1}p \\
&= \sum_{i=1}^{\infty}(i - 1)q^{i-1}p + \sum_{i=1}^{\infty} q^{i-1}p \\
&= \sum_{j=0}^{\infty} jq^{j}p + 1 \\
&= q\sum_{j=1}^{\infty} jq^{j-1}p + 1 \\
&= qE[X] + 1
\end{aligned}$$

Hence,

$$pE[X] = 1$$

yielding the result

$$E[X] = \frac{1}{p}$$

In other words, if independent trials having a common probability p of being successful are performed until the first success occurs, then the expected number of required trials equals 1/p. For instance, the expected number of rolls of a fair die that it takes to obtain the value 1 is 6.

4.8.2 The Negative Binomial Random Variable

Suppose that independent trials, each having probability $p, 0 < p < 1$, of being a success are performed until a total of r successes is accumulated. If we let X equal the number of trials required, then

$$P\{X = n\} = \binom{n-1}{r-1} p^r (1-p)^{n-r} \quad n = r, r+1, \ldots \tag{8.2}$$

Equation (8.2) follows because, in order for the rth success to occur at the nth trial, there must be $r - 1$ successes in the first $n - 1$ trials and the nth trial must be a success. The probability of the first event is

$$\binom{n-1}{r-1} p^{r-1} (1-p)^{n-r}$$

and the probability of the second is p; thus, by independence, Equation (8.2) is established. To verify that a total of r successes must eventually be accumulated, either we can prove analytically that

$$\sum_{n=r}^{\infty} P\{X = n\} = \sum_{n=r}^{\infty} \binom{n-1}{r-1} p^r (1-p)^{n-r} = 1 \tag{8.3}$$

or we can give a probabilistic argument as follows: The number of trials required to obtain r successes can be expressed as $Y_1 + Y_2 + \cdots + Y_r$, where Y_1 equals the number of trials required for the first success, Y_2 the number of additional trials after the first success until the second success occurs, Y_3 the number of additional trials until the third success, and so on. Because the trials are independent and all have the same probability of success, it follows that $Y_1, Y_2, \ldots, Y_r$ are all geometric random variables. Hence, each is finite with probability 1, so $\sum_{i=1}^{r} Y_i$ must also be finite, establishing Equation (8.3).

Any random variable X whose probability mass function is given by Equation (8.2) is said to be a *negative binomial* random variable with parameters (r, p). Note that a geometric random variable is just a negative binomial with parameter $(1, p)$.

In the next example, we use the negative binomial to obtain another solution of the problem of the points.

4.8.3 The Hypergeometric Random Variable

Suppose that a sample of size n is to be chosen randomly (without replacement) from an urn containing N balls, of which m are white and $N - m$ are black. If we let X denote the number of white balls selected, then

$$P\{X = i\} = \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}} \quad i = 0, 1, \ldots, n \tag{8.4}$$

A random variable X whose probability mass function is given by Equation (8.4) for some values of n, N, m is said to be a *hypergeometric* random variable.

Remark. Although we have written the hypergeometric probability mass function with i going from 0 to n, $P\{X = i\}$ will actually be 0, unless i satisfies the inequalities $n - (N - m) \le i \le \min(n, m)$. However, Equation (8.4) is always valid because of our convention that $\binom{r}{k}$ is equal to 0 when either $k < 0$ or $r < k$. ■

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. However, there also exist random variables whose set of possible values is uncountable. Two examples are the time that a train arrives at a specified stop and the lifetime of a transistor. Let X be such a random variable. 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 B of real numbers,‡

$$P\{X \in B\} = \int_B f(x)\,dx \tag{1.1}$$

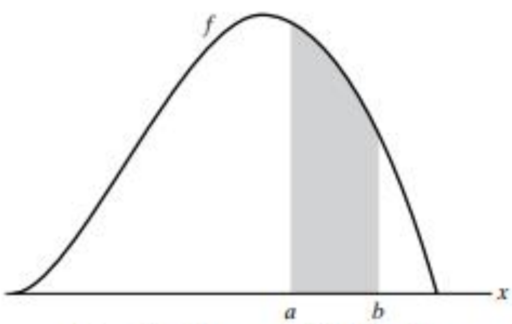
The function f is called the *probability density function* of the random variable X. (See Figure 5.1.)

In words, Equation (1.1) states that the probability that X will be in B may be obtained by integrating the probability density function over the set B. Since X must assume some value, f must satisfy

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

All probability statements about X can be answered in terms of f. For instance, from Equation (1.1), letting $B = [a, b]$, we obtain

$$P\{a \le X \le b\} = \int_a^b f(x)\,dx \tag{1.2}$$



$P(a \le X \le b)$ = area of shaded region

FIGURE 5.1: Probability density function f.

If we let $a = b$ in Equation (1.2), we get

$$P\{X = a\} = \int_a^a f(x)\,dx = 0$$

In words, this equation states that the probability that a continuous random variable will assume any fixed value is zero. Hence, for a continuous random variable,

$$P\{X < a\} = P\{X \le a\} = F(a) = \int_{-\infty}^{a} f(x)\,dx$$

THE UNIFORM RANDOM VARIABLE

A random variable is said to be *uniformly* distributed over the interval (0, 1) if its probability density function is given by

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \tag{3.1}$$

Note that Equation (3.1) is a density function, since $f(x) \geq 0$ and $\int_{-\infty}^{\infty} f(x)\,dx = \int_0^1 dx = 1$. Because $f(x) > 0$ only when $x \in (0,1)$, it follows that X must assume a value in interval (0, 1). Also, since $f(x)$ is constant for $x \in (0,1)$, X is just as likely to

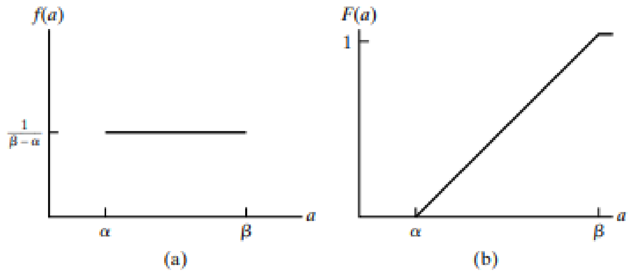


FIGURE 5.3: Graph of (a) $f(a)$ and (b) $F(a)$ for a uniform (α, β) random variable.

be near any value in (0, 1) as it is to be near any other value. To verify this statement, note that, for any $0 < a < b < 1$,

$$P\{a \leq X \leq b\} = \int_a^b f(x)\,dx = b - a$$

In other words, the probability that X is in any particular subinterval of (0, 1) equals the length of that subinterval.

In general, we say that X is a uniform random variable on the interval (α, β) if the probability density function of X is given by

$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha} & \text{if } \alpha < x < \beta \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

Since $F(a) = \int_{-\infty}^{a} f(x)\,dx$, it follows from Equation (3.2) that the distribution function of a uniform random variable on the interval (α, β) is given by

$$F(a) = \begin{cases} 0 & a \leq \alpha \\ \dfrac{a - \alpha}{\beta - \alpha} & \alpha < a < \beta \\ 1 & a \geq \beta \end{cases}$$

Figure 5.3 presents a graph of $f(a)$ and $F(a)$.

EXAMPLE 3a

Let X be uniformly distributed over (α, β). Find (a) $E[X]$ and (b) $\text{Var}(X)$.

Solution. (a)

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f(x)\,dx \\ &= \int_{\alpha}^{\beta} \frac{x}{\beta - \alpha}\,dx \\ &= \frac{\beta^2 - \alpha^2}{2(\beta - \alpha)} \\ &= \frac{\beta + \alpha}{2} \end{aligned}$$

In words, the expected value of a random variable that is uniformly distributed over some interval is equal to the midpoint of that interval.

(b) To find $\mathrm{Var}(X)$, we first calculate $E[X^2]$.

$$E[X^2] = \int_\alpha^\beta \frac{1}{\beta - \alpha} x^2 \, dx = \frac{\beta^3 - \alpha^3}{3(\beta - \alpha)} = \frac{\beta^2 + \alpha\beta + \alpha^2}{3}$$

Hence,

$$\mathrm{Var}(X) = \frac{\beta^2 + \alpha\beta + \alpha^2}{3} - \frac{(\alpha + \beta)^2}{4} = \frac{(\beta - \alpha)^2}{12}$$

Therefore, the variance of a random variable that is uniformly distributed over some interval is the square of the length of that interval divided by 12. ■

Normal Probability Distribution

Bell curves and z-scores

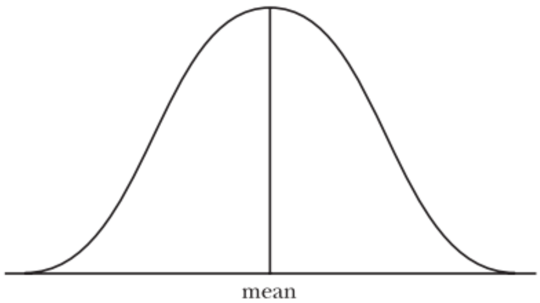
7.1 Identify the three defining characteristics of the normal probability distribution.

A continuous random variable is usually a measurement. Not only can it have integer values, it can also have all the decimal values that fall between integers. See Problem 5.28 for more details.

The normal probability distribution is a bell-shaped continuous distribution that fulfills the following conditions:

- The distribution is symmetrical around the mean.
- The mean, median, and mode are the same value.
- The total area under the curve is equal to one.

The shape of the normal probability distribution is shown below.



Because the normal distribution is continuous, it represents infinitely many possible values, depending on the level of precision. Because there are an infinite number of possible values, the probability that a continuous random variable is equal to a specific single value is zero.

Instead of determining the probability of a single value occurring, when exploring normal distributions, you define two endpoints and calculate the probability that a chosen value will occur within the specified interval.

7.2 Describe the role that the mean, standard deviation, and *z*-score play in the normal probability distribution.

The mean μ is the center of a normal distribution. A higher mean shifts the position of the probability distribution to the right while a lower mean shifts its position to the left.

The standard deviation σ is a measure of dispersion—the higher the standard deviation, the wider the distribution. A smaller standard deviation results in a narrower bell-shaped curve.

The *z*-score measures the number of standard deviations between the mean and a specific value of *x*, according to the formula below.

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

NORMAL RANDOM VARIABLES

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

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

This density function is a bell-shaped curve that is symmetric about μ. (See Figure 5.5.)

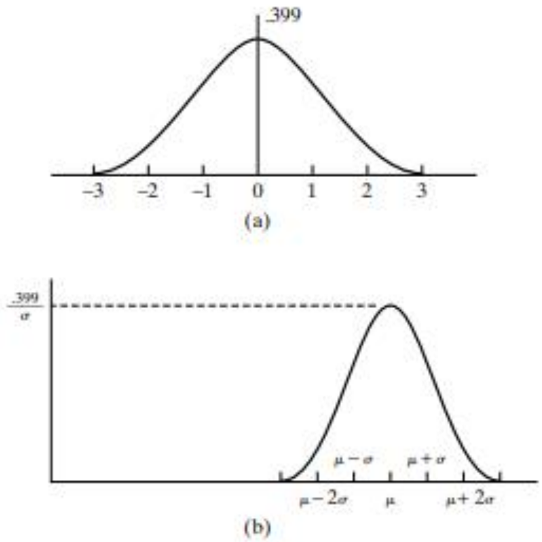


FIGURE 5.5: Normal density function: (a) $\mu = 0, \sigma = 1$; (b) arbitrary μ, σ^2.

The normal distribution was introduced by the French mathematician Abraham DeMoivre in 1733, who used it to approximate probabilities associated with binomial random variables when the binomial parameter n is large. This result was later extended by Laplace and others and is now encompassed in a probability theorem known as the central limit theorem, which is discussed in Chapter 8. The central limit theorem, one of the two most important results in probability theory,† gives a theoretical base to the often noted empirical observation that, in practice, many random phenomena obey, at least approximately, a normal probability distribution. Some examples of random phenomena obeying this behavior are the height of a man, the velocity in any direction of a molecule in gas, and the error made in measuring a physical quantity.

An important implication of the preceding result is that 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 is said to be a standard, or a unit, normal random variable. We now show that the parameters μ and σ^2 of a normal random variable represent, respectively, its expected value and variance.

The Empirical Rule

One, two, and three standard deviations from the mean

7.23 According to the empirical rule, how much of a normally distributed data set lies within one, two, and three standard deviations of the mean?

According to the empirical rule, 68% of the data lies within one standard deviation of the mean, 95% of the data lies within two standard deviations, and 99.7% of the data lies within three standard deviations.

That's one standard deviation above and one below the mean.

EXAMPLE 4a

Find $E[X]$ and $\text{Var}(X)$ when X is a normal random variable with parameters μ and σ^2.

Solution. Let us start by finding the mean and variance of the standard normal random variable $Z = (X - \mu)/\sigma$. We have

$$\begin{aligned} E[Z] &= \int_{-\infty}^{\infty} x f_Z(x)\,dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-x^2/2}\,dx \\ &= -\frac{1}{\sqrt{2\pi}} e^{-x^2/2}\Big|_{-\infty}^{\infty} \\ &= 0 \end{aligned}$$

Thus,

$$\begin{aligned} \text{Var}(Z) &= E[Z^2] \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2}\,dx \end{aligned}$$

Integration by parts (with $u = x$ and $dv = xe^{-x^2/2}$) now gives

$$\begin{aligned} \text{Var}(Z) &= \frac{1}{\sqrt{2\pi}} \left(-xe^{-x^2/2}\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-x^2/2}\,dx\right) \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2}\,dx \\ &= 1 \end{aligned}$$

Because $X = \mu + \sigma Z$, the preceding yields the results

$$E[X] = \mu + \sigma E[Z] = \mu$$

and

$$\text{Var}(X) = \sigma^2 \text{Var}(Z) = \sigma^2 \qquad \blacksquare$$

Historical Notes Concerning the Normal Distribution

The normal distribution was introduced by the French mathematician Abraham DeMoivre in 1733. DeMoivre, who used this distribution to approximate probabilities connected with coin tossing, called it the exponential bell-shaped curve.

Its usefulness, however, became truly apparent only in 1809, when the famous German mathematician Karl Friedrich Gauss used it as an integral part of his approach to predicting the location of astronomical entities. As a result, it became common after this time to call it the Gaussian distribution.

During the mid- to late 19th century, however, most statisticians started to believe that the majority of data sets would have histograms conforming to the Gaussian bell-shaped form. Indeed, it came to be accepted that it was “normal” for any well-behaved data set to follow this curve. As a result, following the lead of the British statistician Karl Pearson, people began referring to the Gaussian curve by calling it simply the normal curve. (A partial explanation as to why so many data sets conform to the normal curve is provided by the central limit theorem, which is presented in Chapter 8.)

Abraham DeMoivre (1667–1754)

Today there is no shortage of statistical consultants, many of whom ply their trade in the most elegant of settings. However, the first of their breed worked, in the early years of the 18th century, out of a dark, grubby betting shop in Long Acres, London, known as Slaughter's Coffee House. He was Abraham DeMoivre, a Protestant refugee from Catholic France, and, for a price, he would compute the probability of gambling bets in all types of games of chance. Although DeMoivre, the discoverer of the normal curve, made his living at the coffee shop, he was a mathematician of recognized abilities. Indeed, he was a member of the Royal Society and was reported to be an intimate of Isaac Newton. Listen to Karl Pearson imagining DeMoivre at work at Slaughter's Coffee House: "I picture DeMoivre working at a dirty table in the coffee house with a brokendown gambler beside him and Isaac Newton walking through the crowd to his corner to fetch out his friend. It would make a great picture for an inspired artist."

Karl Friedrich Gauss (1777–1855), one of the earliest users of the normal curve, was one of the greatest mathematicians of all time. Listen to the words of the well-known mathematical historian E. T. Bell, as expressed in his 1954 book Men of Mathematics: In a chapter entitled "The Prince of Mathematicians," he writes, "Archimedes, Newton, and Gauss; these three are in a class by themselves among the great mathematicians, and it is not for ordinary mortals to attempt to rank them in order of merit. All three started tidal waves in both pure and applied mathematics. Archimedes esteemed his pure mathematics more highly than its applications; Newton appears to have found the chief justification for his mathematical inventions in the scientific uses to which he put them; while Gauss declared it was all one to him whether he worked on the pure or on the applied side."

5.4.1 The Normal Approximation to the Binomial Distribution

An important result in probability theory known as the DeMoivre–Laplace limit theorem states that when n is large, a binomial random variable with parameters n and p will have approximately the same distribution as a normal random variable with the same mean and variance as the binomial. This result was proved originally for the special case of $p = \frac{1}{2}$ by DeMoivre in 1733 and was then extended to general p by Laplace in 1812. It formally states that if we "standardize" the binomial by first subtracting its mean np and then dividing the result by its standard deviation $\sqrt{np(1 - p)}$, then the distribution function of this standardized random variable (which has mean 0 and variance 1) will converge to the standard normal distribution function as $n \to \infty$.

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 $a < b$,

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

as $n \to \infty$.

Because the preceding theorem is only a special case of the central limit theorem, which is presented in Chapter 8, we shall not present a proof.

Note that we now have two possible approximations to binomial probabilities: the Poisson approximation, which is good when n is large and p is small, and the normal approximation, which can be shown to be quite good when $np(1 - p)$ is large. (See Figure 5.6.) [The normal approximation will, in general, be quite good for values of n satisfying $np(1 - p) \geq 10$.]

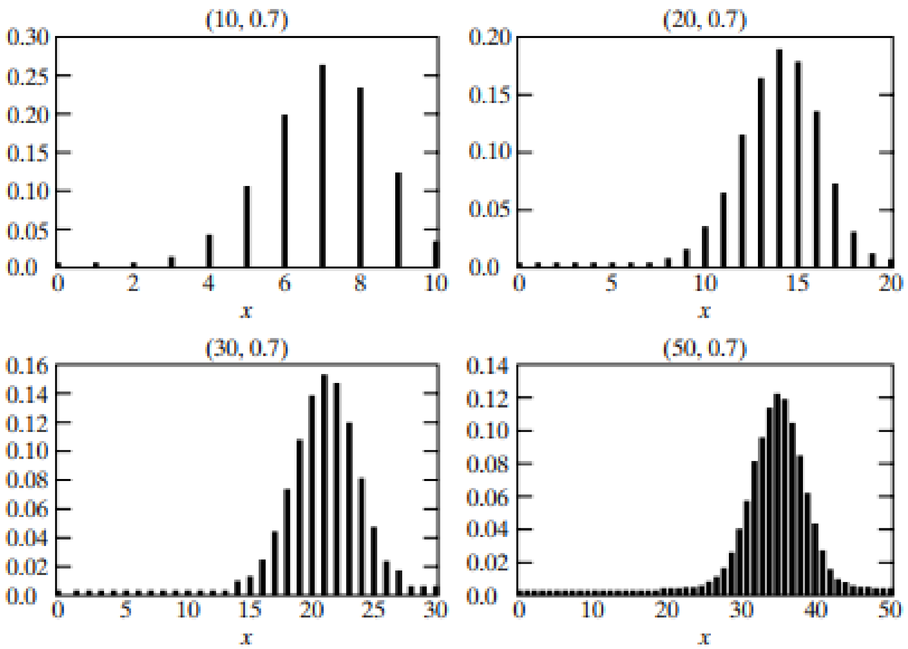


FIGURE 5.6: The probability mass function of a binomial (n, p) random variable becomes more and more "normal" as n becomes larger and larger.

Note: In Problems 7.19–7.22, assume that the number of days it takes a homebuilder to complete a house is normally distributed with an average completion time of 176.7 days and a standard deviation of 24.8 days.

7.19 Calculate the probability that it will take between 185 and 225 days to complete the next home.

Calculate the z-scores for $x = 185$ and $x = 225$.

$$z_{185} = \frac{185-176.7}{24.8} = \frac{8.3}{24.8} = 0.33 \qquad z_{225} = \frac{225-176.7}{24.8} = \frac{48.3}{24.8} = 1.95$$

Note that $P(0 < z < 0.33) = 0.1293$ and $P(0 < z < 1.95) = 0.4744$.

$$\begin{aligned} P(185 < x < 225) &= P(0.33 < z < 1.95) \\ &= 0.4744 - 0.1293 \\ &= 0.3451 \end{aligned}$$

Using the Normal Distribution to Approximate the Binomial Distribution

Another binomial probability shortcut

7.30 Describe the conditions under which the normal distribution can be used to approximate the binomial distribution.

When it comes to binomial distributions, p is the probability of a success and q is the probability of a failure. "Success" means you got the one result out of the two you were looking for.

If n represents the number of trials in which only outcomes p and q may occur, the normal distribution can be used to approximate the binomial distribution as long as $np \geq 5$ and $nq \geq 5$.

7.31 Describe the continuity correction that is applied when the normal distribution approximates the binomial distribution.

Continuity correction is used when a continuous distribution (such as the normal distribution) is used to approximate a discrete distribution (such as the binomial distribution). To correct for continuity, add 0.5 to a boundary of x or subtract 0.5 from a boundary of x as directed below:

- Subtract 0.5 from the x-value representing the left boundary under the normal curve.
- Add 0.5 to the x-value representing the right boundary under the normal curve.

Note that continuity correction is unnecessary when $n > 100$.

Note: Problems 7.32–7.33 refer to a statistics class in which 60% of the students are female; 15 students from the class are randomly selected.

The problem doesn't say how big the class is, but it's got to be bigger than 15 students if you're selecting that many of them randomly.

7.32 Use the normal approximation to the binomial distribution to calculate the probability that this randomly selected group will contain either seven or eight female students.

Determine whether conditions have been met to use the normal distribution to approximate the binomial distribution.

$$np = (15)(0.6) = 9;\ 9 \geq 5$$
$$nq = (15)(0.4) = 6;\ 6 \geq 5$$

These formulas come from Problem 6.5.

Calculate the mean and standard deviation of the binomial distribution.

$$\mu = np = (15)(0.6) = 9$$
$$\sigma = \sqrt{npq} = \sqrt{(15)(0.6)(0.4)} = \sqrt{3.6} = 1.90$$

The problem asks you to calculate $P(7 \leq x \leq 8)$. Apply the continuity correction to adjust the boundaries: $P(6.5 \leq x \leq 8.5)$.

Subtract 0.5 from the left boundary x = 7 and add 0.5 to the right boundary x = 8.

Calculate the z-scores for endpoints $x = 6.5$ and 8.5.

$$z_{6.5} = \frac{6.5 - 9}{1.90} = \frac{-2.5}{1.90} = -1.32 \qquad z_{8.5} = \frac{8.5 - 9}{1.90} = \frac{-0.5}{1.90} = -0.26$$

According to Reference Table 1, $P(-1.32 < z < 0) = 0.4066$ and $P(-0.26 < z < 0) = 0.1026$.

$$\begin{aligned} P(6.5 \le x \le 8.5) &= P(-1.32 \le z \le -0.26) \\ &= 0.4066 - 0.1026 \\ &= 0.3040 \end{aligned}$$

> When both x values are on the same side of the mean (in this case, they're both below 9) and you're calculating the area of the region between them, subtract the smaller probability from the larger probability.

There is a 30.4% chance that the group of 15 students will contain either seven or eight females.

EXPONENTIAL RANDOM VARIABLES

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

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

is said to be an *exponential* random variable (or, more simply, is said to be exponentially distributed) with parameter λ. The cumulative distribution function $F(a)$ of an exponential random variable is given by

$$\begin{aligned} F(a) &= P\{X \le a\} \\ &= \int_0^a \lambda e^{-\lambda x}\, dx \\ &= -e^{-\lambda x}\Big|_0^a \\ &= 1 - e^{-\lambda a} \qquad a \ge 0 \end{aligned}$$

Note that $F(\infty) = \int_0^\infty \lambda e^{-\lambda x}\, dx = 1$, as, of course, it must. The parameter λ will now be shown to equal the reciprocal of the expected value.

EXAMPLE 5a

Let X be an exponential random variable with parameter λ. Calculate (a) $E[X]$ and (b) $\mathrm{Var}(X)$.

Solution. (a) Since the density function is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & x < 0 \end{cases}$$

we obtain, for $n > 0$,

$$E[X^n] = \int_0^\infty x^n \lambda e^{-\lambda x}\, dx$$

Integrating by parts (with $\lambda e^{-\lambda x} = dv$ and $u = x^n$) yields

$$\begin{aligned} E[X^n] &= -x^n e^{-\lambda x}\Big|_0^\infty + \int_0^\infty e^{-\lambda x} n x^{n-1}\, dx \\ &= 0 + \frac{n}{\lambda}\int_0^\infty \lambda e^{-\lambda x} x^{n-1}\, dx \\ &= \frac{n}{\lambda} E[X^{n-1}] \end{aligned}$$

Letting $n = 1$ and then $n = 2$ gives

$$E[X] = \frac{1}{\lambda}$$

$$E[X^2] = \frac{2}{\lambda}E[X] = \frac{2}{\lambda^2}$$

(b) Hence,

$$\text{Var}(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

Thus, the mean of the exponential is the reciprocal of its parameter λ, and the variance is the mean squared. ■

In practice, the exponential distribution often arises as the distribution of the amount of time until some specific event occurs. For instance, the amount of time (starting from now) until an earthquake occurs, or until a new war breaks out, or until a telephone call you receive turns out to be a wrong number are all random variables that tend in practice to have exponential distributions. (For a theoretical explanation of this phenomenon, see Section 4.7.)

Exponential Distribution

Like the Poisson distribution, but continuous

7.45 Identify the defining characteristics of the exponential distribution.

The exponential probability distribution is a continuous distribution commonly used to measure the time between events of interest, such as the time between customer arrivals at a retail store or the time between failures in a process.

In Chapter 6, the variable λ was used to represent the mean of the Poisson distribution, a discrete distribution that counted the number of times an event occurred during a specific time period. The mean and the standard deviation of the exponential distribution are both $\frac{1}{\lambda}$.

The exponential distribution is the continuous counterpart of the discrete Poisson distribution. For example, if a random variable follows the Poisson distribution with an average occurrence of two times per minute ($\lambda = 2$), then the same random variable also follows the exponential distribution with a mean and standard deviation of $\frac{1}{\lambda} = \frac{1}{2} = 0.5$.

If x is a random variable that follows the exponential distribution, then the probability that $x \geq t$ is $e^{-\lambda t}$.

$$P(x \geq t) = e^{-\lambda t}$$

The constant e is Euler's number, the nonrepeating, nonterminating decimal 2.71828....

Note: In Problems 7.49–7.51, assume that the tread life of a particular brand of tire is exponentially distributed and averages 32,000 miles.

7.49 Calculate the probability that a set of these tires will have a tread life of at least 38,000 miles.

The mean of the distribution is $\frac{1}{\lambda} = 32$ miles per set of tires (in thousands), so $\lambda = \frac{1}{32} = 0.03125$ sets of tires per thousand miles. Substitute $t = 38$ into the exponential probability formula to calculate the probability that a particular set of tires will have a tread life of more than 38,000 miles.

$$P(x \geq t) = e^{-\lambda t}$$
$$P(x \geq 38) = e^{-(0.03125)(38)}$$
$$P(x \geq 38) = e^{-1.1875}$$
$$P(x \geq 38) = 0.305$$

Chapter 8

SAMPLING AND SAMPLING DISTRIBUTIONS

Working with a subset of a population

A population is defined as all possible outcomes or measurements of interest, whereas a sample is a subset of a population. Many populations are infinitely large; thus, virtually all statistical analyses are conducted on samples drawn from a population. In order to interpret the results of these analyses correctly, you must first understand the behavior of samples. In this chapter, you will do just that through the exploration of sampling distributions.

This chapter relies heavily on the normal probability distribution concepts introduced in Chapter 7. The two major topics are the sampling distribution of the mean and the sampling distribution of the proportion. Also make sure you understand binomial distributions, as they make a guest appearance late in the chapter.

Probability Sampling

So many ways to gather a sample

8.1 Describe how to select a simple random sample from a population.

A simple random sample is a sample that is randomly selected so that every combination has an equal chance of being chosen. If an urn contains six balls of different colors, selecting three of the balls without looking inside the urn is an example of a simple random sample.

8.2 Describe how to select a systematic sample from a population.

Systematic sampling includes every *k*th member of the population in the sample; the value of *k* will depend on the size of the population and the size of the sample that is desired. For instance, if a sample size of 50 is needed from a population of 1,000, then $k = \frac{1{,}000}{50} = 20$. Systematically, every twentieth person from the population is selected and included in the sample.

8.3 Describe how to select a cluster sample from a population.

Cluster sampling first divides the population into groups (or clusters) and then randomly selects clusters to include in the sample. The entire cluster or just a randomly selected portion of it may be selected. For example, if a researcher wishes to poll a sample of shoppers at a shopping mall, she might choose a few stores randomly, and then interview the customers inside those stores only. In this example, the stores are the clusters.

In order for cluster sampling to be effective, each cluster selected for the sample needs to be representative of the population at large.

> Cluster sampling is cost-effective because it requires minimal research about the population. In the mall example, you didn't have to know anything about the shoppers ahead of time—you just needed to pick a few stores from the map.

8.4 Describe how to select a stratified sample from a population.

Stratified sampling first divides the population into mutually exclusive groups (or strata) and then selects a random sample from each of those groups. It differs from cluster sampling in that strata are defined in terms of specific characteristics of the population, whereas clusters produce less homogeneous samples.

Consider the example presented in Problem 8.3, in which clusters are assigned based upon the stores in a mall. A stratified sample would be chosen in terms of a specific customer characteristic, such as gender. Stratified sampling is helpful when it is important that the sample have certain characteristics of the overall population. Usually the sample sizes are proportional to their known relationship in the population.

> If cluster sampling had been used at the mall to ask how male teenagers respond to a new product, there's no guarantee that the cluster sample would have included male teenagers at all.

Sampling Distribution of the Mean

Predicting the behavior of sample means

8.5 Identify the implications of the central limit theorem on the sampling distribution of the mean.

> It's called the CENTRAL limit theorem because it's the most important theorem in statistics.

According to the central limit theorem, as a sample size n gets larger, the distribution of the sample means more closely approximates a normal distribution, regardless of the distribution of the population from which the sample was drawn. As a general rule of thumb, the assertions of the central limit theorem are valid when $n \geq 30$. If the population itself is normally distributed, the sampling distribution of the mean is normal for any sample size.

As the sample size increases, the distribution of sample means converges toward the center of the distribution. Thus, as the sample size increases, the standard deviation of the sample means decreases. According to the central limit theorem, the standard deviation of the sample means $\sigma_{\bar{x}}$ is equal to $\frac{\sigma}{\sqrt{n}}$, where σ is the standard deviation of the population and n is the sample size.

The standard deviation of the sample mean is formally known as the standard error of the mean. The z–score for sample means is calculated based on the formula below.

$$z_{\bar{x}} = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

> The variable $\bar{x}$ represents the mean of the sample.

Note: In Problems 8.6–8.8, assume that the systolic blood pressure of 30-year-old males is normally distributed, with an average of 122 mmHg and a standard deviation of 10 mmHg.

The unit mmHg stands for "millimeters of mercury."

8.6 A random sample of 16 men from this age group is selected. Calculate the probability that the average blood pressure of the sample will be greater than 125 mmHg.

The population is normally distributed, so sample means are also normally distributed for any sample size. Calculate the standard error of the mean.

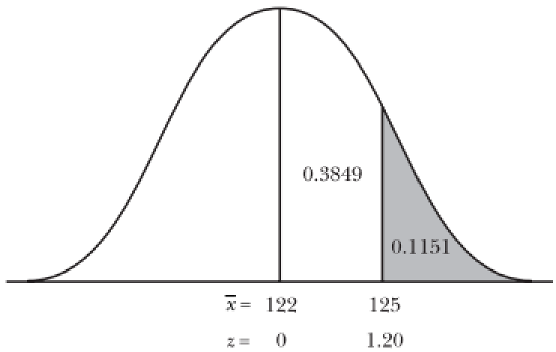
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$
$$\sigma_{125} = \frac{10}{\sqrt{16}}$$
$$\sigma_{125} = \frac{10}{4}$$
$$\sigma_{125} = 2.5$$

Calculate the z-score for the sample mean, $\bar{x} = 125$.

$$z_{\bar{x}} = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$
$$z_{125} = \frac{125 - 122}{2.5}$$
$$z_{125} = \frac{3}{2.5}$$
$$z_{125} = 1.2$$

See Problem 7.3 if you're not sure how to use Reference Table 1.

According to Reference Table 1, the normal probability associated with $z = 1.20$ is 0.3849. The probability that the sample mean will be greater than 125 is the area of the shaded region beneath the normal curve in the figure below. The area below the curve on each side of the mean is 0.5, and the area between the mean and the z-score 1.20 is 0.3849.



Calculate the probability that the average blood pressure of the sample will be greater than 125 mmHg.

$$P(\bar{x} > 125) = P(z_{\bar{x}} > 1.20)$$
$$= 0.5 - 0.3849$$
$$= 0.1151$$

Finite Population Correction Factor

Sampling distribution of the mean with a small population

8.18 Describe the finite population correction factor for the sampling distribution for the mean and the conditions under which it should be applied.

When a population is very large, selecting something as part of a sample has a negligible impact on the population. For instance, if you randomly chose individuals from the continent of Europe and recorded the gender of the individuals you chose, selecting a finite number of men would not significantly change the probability that the next individual you chose would also be male.

However, when the sample size n is more than 5 percent of the population size N, the finite population correction factor below should be applied. Under this condition, the population size is small enough that the sampling events are no longer independent of one another. The selection of one item from the population impacts the probability of future items being selected.

In other words, when $\frac{n}{N} > 0.05$. Some textbooks say 10% instead of 5%.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

Note: Problems 8.19–8.20 refer to a process that fills boxes with a mean of 340 grams of cereal, with a standard deviation of 20 grams. Assume the probability distribution for this population is unknown.

8.19 If a store purchases 600 boxes of cereal, what is the probability that a sample of 50 boxes from the order will average less than 336 grams?

Note that the sample is more than 5% of the total population: $\frac{n}{N} = \frac{50}{600} = 0.083 > 0.05$. Thus, you must apply the finite population correction factor when calculating the standard error of the mean.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{20}{\sqrt{50}}\sqrt{\frac{600-50}{600-1}} = \frac{20}{\sqrt{50}}\sqrt{\frac{550}{599}} = 2.710$$

Without the correction, the standard error of the mean is 2.83.

Calculate the z-score for $\bar{x} = 336$.

$$z_{\bar{x}} = \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} = \frac{336-340}{2.710} = \frac{-4}{2.710} = -1.48$$

There is a 0.5 probability that the sample mean will be less than the population mean of 340. According to Reference Table 1, there is a 0.4306 probability that the sample mean will be between 336 and 340.

$$\begin{aligned} P(\bar{x} < 336) &= P(z_{\bar{x}} < -1.48) \\ &= P(z_{\bar{x}} < 0) - P(-1.48 < z_{\bar{x}} < 0) \\ &= 0.50 - 0.4306 \\ &= 0.0694 \end{aligned}$$

Sampling Distribution of the Proportion

Predicting the behavior of discrete random variables

8.23 Describe the sampling distribution of the proportion and the circumstances under which it is used.

The sampling distribution of the proportion is applied when the random variable is binomially distributed. Divide the number of successes s by the sample size n to calculate p_s, the proportion of successes in the sample.

$$p_s = \frac{s}{n}$$

Calculate the standard error of the proportion σ_p by substituting the population proportion p (not the sample proportion p_s) into the formula below.

$$\sigma_p = \sqrt{\frac{p(1-p)}{n}}$$

Make sure that p is between zero and one. If p is greater than one, 1 − p will be negative and your calculator will explode when you try to take the square root of a negative number.

Yararlanılan Kaynaklar

A First Course in Probability (Sheldon Ross, 8. Baskı, 2010)

Hand Book of Statistical Terms (Nobel Yayın Dağıtım,2010)

The Humongous Books of Statistics Problems

https://www.oxfordreference.com/view/10.1093/acref/9780199541454.001.0001/acref-9780199541454

https://www.stat.berkeley.edu/~stark/SticiGui/Text/gloss.htm