BUSAD40 - Busniess Statistics

Lecture Note 2

Norman Lo

Spring 2021

It has been pointed out already that no knowledge of probabilities, less in degree than certainty, helps us to know what conclusions are true, and that there is no direct relation between the truth of a proposition and its probability. Probability begins and ends with probability.

— John Maynard Keynes, The Application of Probability to Conduct

I. Introduction to Probability

Managers often base their decisions on an analysis of uncertainties such as the following:

- What are the chances that sales will decrease if we increase prices?
- What is the likelihood a new assembly method will increase productivity?
- What are the odds that a new investment will be profitable?

Probability is a numerical measure of the likelihood that an event will occur. Probability values are always assigned on a scale from 0 to 1. A probability near zero indicates an event is quite unlikely to occur and near one indicates an event is almost certain to occur.

The idea of probability as a proportion of outcome in very many repeated trials guides our intuition but is hard to express in mathematical form. A description of a random phenomenon in the language of mathematics is called a **Probability Model**.

Statistical Experiments:

In statistics, the notion of an experiment differs somewhat from that of an experiment in the physical sciences. In statistical experiments, probability determines outcomes. Even though the experiment is repeated in the same way, an entirely different outcome may occur. For this reason, statistical experiments are sometimes called **random experiments**. A Random experiment is a process that generates well-defined experimental outcomes. The **sample space** for an experiment is the set of all experimental outcomes. An experimental outcome is also called a **sample point**.

Experiment	Experiment Outcomes
Toss a coin	Head, Tail
Inspect a part	Defective, Non-Defective
Conduct a sales call	Purchase, No Purchase
Roll a die	1, 2, 3, 4, 5, 6
Play a football game	Win, Lose, Tie

Counting Rule for Multiple-Step Experiments

If an experiment consists of a sequence of k steps in which there are n_1 possible results for the first step, n_2 possible results for the second step, and so on, then the total number of experimental outcomes is given by (n_1)(n_2) . . . (n_k).

Example:

An investor is planning to invest on two stocks. Each stock has a probability to gain, lose, or tie. The total combination of the outcomes from investing to these two stocks are 3 outcomes from stock A and 3 outcomes from stock B, total of 3 x 3 = 9 total combination outcomes.

Assigning Probabilities

Basic Requirements for Assigning Probabilities,

1. The probability assigned to each experimental outcome must be between 0 and 1, inclusively.
2. The sum of the probabilities for all experimental outcomes must equal 1.

Three common methods to assign probability to an event,

1. Classical Method: Assigning probabilities based on the assumption of equally likely outcomes
2. Relative Frequency Method: Assigning probabilities based on experimentation or historical data
3. Subjective Method: Assigning probabilities based on judgment

Events and Their Probabilities

An event is a collection of sample points. The probability of any event is equal to the sum of the probabilities of the sample points in the event. If we can identify all the sample points of an experiment and assign a probability to each, we can compute the probability of an event.

Basic Relationship of Probability

There are some basic probability relationships that can be used to compute the probability of an event without knowledge of all the sample point probabilities.

Complement of an Event

The complement of event A is defined to be the event consisting of all sample points that are not in A.

e.g. Suppose event A is getting a number 3 from rolling a fair dice. Complement of event A is all the possible outcome that is not number 3.

Probability of Event A = P(A) = 1/6

Complement of Event A = P(Not A) = 1 - 1/6 = 5/6

Union of Two Event

The union of events A and B is the event containing all sample points that are in A or B or both. The union of events A and B is denoted by $A \bigcup B$.

e.g. Suppose event A is getting a number 3 and event B is getting a number 4 or 6 from rolling a fair dice. The union of event A and B is to roll some dice to get either 3 or 4 and 6.

Probability of Event A = P(A) = 1/6

Probability of Event B = P(B) = 2/6

Probability of Event A union Event B = $P(A\bigcup B)$ = 1/6 + 2/6 = 3/6 = 1/2

Intersection of Two Events

The intersection of events A and B is the set of all sample points that are in both A and B. The intersection of events A and B is denoted by $A\bigcap B$.

e.g. Suppose event A is getting an even number and event B is getting a number greater than 2 from rolling a fair dice. The intersection of event A and B is to roll some dice to get an even number that is greater than 2.

Probability of Event A = P(A) = P(2) + P(4) + P(6)

Probability of Event B = P(B) = P(3) + P(4) + P(5) + P(6)

Intersection of Event A and B = $P(A\bigcap B)$ = P(4) + P(6) = 1/6 + 1/6 = 2/6 = 1/3

Addition Law

The addition law provides a way to compute the probability of event A, or B, or both A and B occurring. The law is written as:

$$P(A\bigcup B) = P(A) + P(B) - P(A\bigcap B)$$

Mutually Exclusive Events

Two events are said to be mutually exclusive if the events have no sample points in common. Two events are mutually exclusive if, when one event occurs, the other cannot occur. If events A and B are mutually exclusive, $P(A\bigcap B) = 0$. The addition law for mutually exclusive events is:

$$P(A\bigcup B) = P(A) + P(B)$$

Conditional Probability

The probability of an event given that another event has occurred is called a conditional probability. The conditional probability of A given B has already occurred is denoted by P(A|B). A conditional probability is computed as follows :

$$P(A|B) = \frac{P(A\bigcap B)}{P(B)}$$

e.g. Suppose we are going to roll a fair dice two times. We want to know the probability for getting a number greater than 2 given the first trial is an even number. The conditional probability of event B given the event A.

Probability of the Intersection of Event A and B = $P(A\bigcap B)$ = 1/3 Probability of Event B = P(B) = 4/6
Condition Probability = P(B|A) = $\frac{P(A\bigcap B)}{P(B)}$ = 1/3 x 6/4 = 2/4 = 1/2

Multiplication Law

The multiplication law provides a way to compute the probability of the intersection of two events. The law is written as:

$$P(A\bigcap B) = P(B)P(A|B)$$

OR

$$P(A\bigcap B) = P(A)P(B|A)$$

Independent Event

If the probability of event A is not changed by the existence of event B, we would say that events A and B are **independent**. Two events A and B are independent if:

$$P(A|B) = P(A)$$

OR

$$P(B|A) = P(B)$$

Multiplication Law for Independent Events

The multiplication law also can be used as a test to see if two events are independent. The law is written as:

$$P(A \bigcap B) = P(A)P(B)$$

Mutually Exclusive and Independent

Do not confuse the notion of mutually exclusive events with that of independent events. Two events with nonzero probabilities cannot be both mutually exclusive and independent. If one mutually exclusive event is known to occur, the other cannot occur.; thus, the probability of the other event occurring is reduced to zero (and they are therefore dependent). Two events that are not mutually exclusive, might or might not be independent.

Examples:

Here is the frequency distribution of College Students classified by age and full-time or part-time status:

Age (years)	Full-time	Part-time	Total:
15 - 19	210	20	230
20 - 24	320	70	390
25 - 29	100	100	200
30 and over	50	130	180
Total:	680	320	1000

1. What is the probability to randomly select a student who is 18 years old?

Rephrase: $P(Age = 18)$ where selecting a student who is 18 years old is Event A.

Ans: $P(Age = 18) = \frac{230}{1000} = 0.23 \; or \; 23\%$

2. What is the probability to randomly select a student who is 22 years old and studying full time?

Rephrase: $P(Age = 22 \cap Status = Part\ time)$ where selecting a student who is 22 years old is Event A and part-time status is Event B.

Ans: $P(Age = 22 \cap Status = Part\ time) = \frac{70}{1000} = 0.07 \; or \; 7\%$

3. What is the probability to randomly select a student who is either over 30 years old or studying part time?

Rephrase: $P(Age > 30 \cup Status = Full\ time)$ where selecting a student who is over 30 years old is Event A and full-time status is Event B.

Ans: $P(Age > 30 \cup Status = Full\ time) = \frac{180+680}{1000} = 0.86\ or\ 86\%$

4. What is the probability to randomly select a student who is 28 years old given that the student is studying part time?

Rephrase: $P(Age = 28|Status = Part\ time)$ where selecting a student who is 28 years hold is Event A and selecting a student who is studying part time is Event B.

Ans: $P(Age = 28|Status = Part\ time) = \frac{\frac{100}{1000}}{\frac{320}{1000}} = \frac{0.1}{0.32} = 0.03125\ or\ 3.125\%$

Bayes' Theorem

Often, we begin probability analysis with initial or **prior probabilities**. Then, from a sample, special report, or a product test we obtain some additional information. Given this information, we calculate revised or **posterior probabilities**. **Bayes' Theorem** provides the means of revising the prior probabilities.

Example:

Suppose you saw some smoke coming from a backyard of the house and you are thinking if the house is caught on fire. Given you some probability of different events, you can apply the Bayes' Theorem to calculate the probability of a house catching fire given that there is smoke coming from the backyard.

Define different probabilities:

Probability a house catches fire = P(fire) = 1%

Probability seeing smoke coming from a backyard = P(smoke) = 10%

Probability seeing smoke given the house is on fire = P(smoke|fire) = 90%

Probability a House catches fire given seeing smoke = P(fire|smoke) = ?

Solution:

$P(fire|smoke) = \frac{P(fire)P(smoke|fire)}{P(smoke)}$

$P(fire|smoke) = \frac{0.01*0.9}{0.1}$

$P(fire|smoke) = 0.09$ OR 9%

II. Probability and Distribution

In the previous section, we demonstrate how to calculate the probability for a given event, a sequence of events, or events by conditions. In this section, we would like to discover how we can apply the probability concepts to different types of data and visualize it on a distribution.

A probability distribution is a function that describes the likelihood of obtaining the possible values that a random variable can assume. In other words, the values of the variable vary based on the underlying probability distribution.

Suppose you draw a random sample and measure the heights of the subjects. As you measure heights, you can create a distribution of heights. This type of distribution is useful when you need to know which outcomes are most likely, the spread of potential values, and the likelihood of different results.

Discrete Probability Distribution

A **discrete probability function** is a function that can take a discrete number of values (not necessarily finite). This is most often the non-negative integers or some subset of the non-negative integers. There is no mathematical restriction that discrete probability functions only be defined at integers, but in practice this is usually what makes sense. For example, if you toss a coin 6 times, you can get 2 heads or 3 heads but not 2 1/2 heads. Each of the discrete values has a certain probability of occurrence that is between zero and one. That is, a discrete function that allows negative values or values greater than one is not a probability function. The condition that the probabilities sum to one means that at least one of the values must occur.

Discrete probability functions are referred to as probability mass functions and continuous probability functions are referred to as probability density functions. The term probability functions cover both discrete and continuous distributions. Here is some discrete probability function you may study in the future:

Binomial Distribution: The experiment consists of a sequence of "**n**" identical trials. Two outcomes, success or failure, are possible on each trial. The probability of a success, denoted by p, does not change from trial to trial. (This is referred to as the stationarity assumption.) The trials are independent.

Binomial Probability Function:

$$f(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{(n-x)}$$

where:

x = the number of successes

p = the probability of a success on one trial

n = the number of trials

f(x) = the probability of x success in n trails

n! = n(n-1)(n-2)…(2)(1)

Binomial Probability Distribution Properties:

Expected Value: E(x) = μ = np

Variance: Var(x) = σ^2 = np(1 - p)

Standard Deviation: $\sigma = \sqrt{np(1-p)}$

Example:

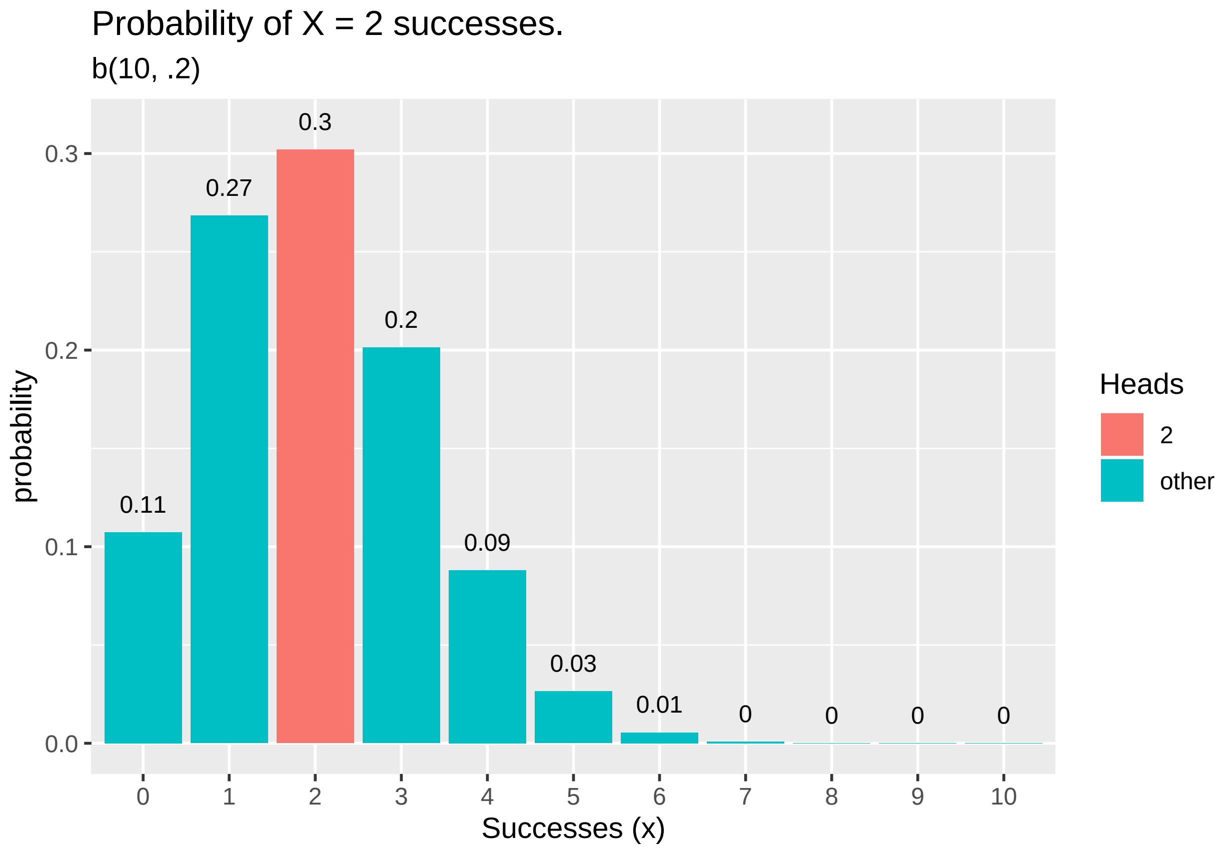
What is the probability of 2 heads in 10-coin flips where probability of heads is 0.2?

```
# Calculate the probability with the given conditions
dbinom(x = 2, size = 10, prob = 0.2)
```

```
## [1] 0.3019899
```

```
# Generate a simulated data
mean(rbinom(n = 10000, size = 10, prob = 0.2) == 2)
```

```
## [1] 0.3019
```



Poisson Probability Distribution: A Poisson distributed random variable is often useful in estimating the number of occurrences over a specified interval of time or space. It is a discrete random variable that may assume an infinite sequence of values (x = 0, 1, 2, . . .). There are two properties of a Poisson experiment, 1) the probability of an occurrence is the same for any two intervals of equal length and 2) the occurrence or nonoccurrence in any interval is independent of the occurrence or nonoccurrence in any other interval.

Poisson Probability Function:

$$f(x) = \frac{\mu^2 e^{-\mu}}{x!}$$

where:

x = the number of occurrences in an interval

f(x) = the probability of x occurrences in an interval

μ = mean number of occurrences in an interval

e = 2.71828. . .

x! = x(x-1)(x-2). . . (2)(1)

Poisson Probability Distribution Properties:
A property of the Poisson distribution is that the mean and variance are equal.

$$\mu = \sigma^2$$

Continuous Probability Distribution

A continuous random variable can assume any value in an interval on the real line or in a collection of intervals. It is not possible to talk about the probability of the random variable assuming a particular value. Instead, we talk about the probability of the random variable assuming a value within a given interval.

The probability of the random variable assuming a value within some given interval from x_1 to x_2 is defined to be the area under the graph of the probability density function between x_1 and x_2.

Uniform Probability Distribution

A random variable is uniformly distributed whenever the probability is proportional to the interval's length. The uniform probability density function is:

$$f(x) = \frac{1}{(b-a)} \; for \; (a \leq x \leq b)$$

$$= 0 \; otherwise$$

where:

a = smallest value the variable can assume

b = largest value the variable can assume

Uniform Probability Distribution Property:

Expected Value of X: $E(x) = \frac{(a+b)}{2}$

Variance of x: $Var(x) = \frac{(b-a)^2}{12}$

Example:

Slater's customers are charged for salad they take. Sampling suggests that the amount of salad taken is uniformly distributed between 5 ounces and 15 ounces.

Uniform Probability Density Function:

$$f(x) = \frac{1}{10} \; for \; (5 \leq x \leq 15)$$

$$= 0 \; otherwise$$

where: x = salad plate filling weight

Expected Value of x: $E(x) = \frac{(5+15)}{2} = 10$

Variance of x: $var(x) = \frac{(15-5)^2}{12} = 8.33$

Normal Probability Distribution

The normal probability distribution is the most important distribution for describing a continuous random variable. It is widely used in statistical inference. It has been used in a wide variety of applications including height of people, test scores, amounts of rainfall, scientific measurements, etc. Abraham de Moivre, a French mathematician, published The Doctrine of Chances in 1733, derived the normal distribution.

Normal Probability Density Function:

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

where:

μ = mean

σ = standard deviation

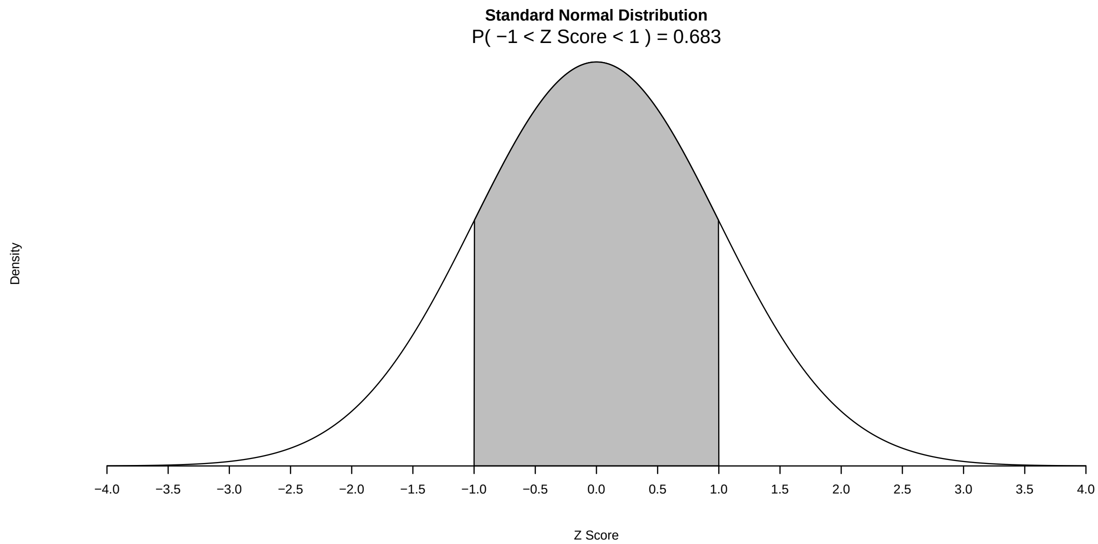
$\pi = 3.14159\ldots$

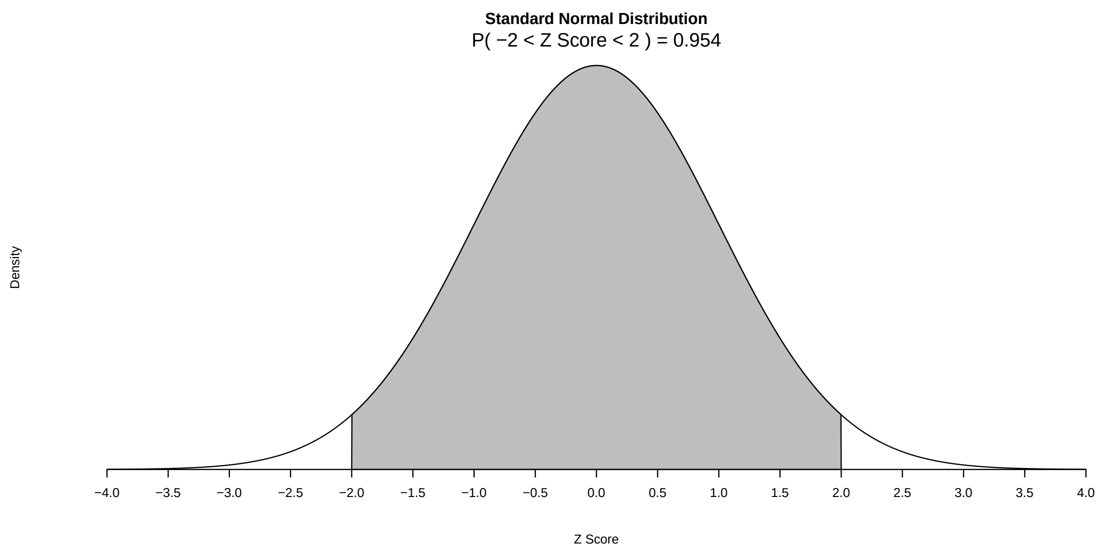
$e = 2.71828\ldots$

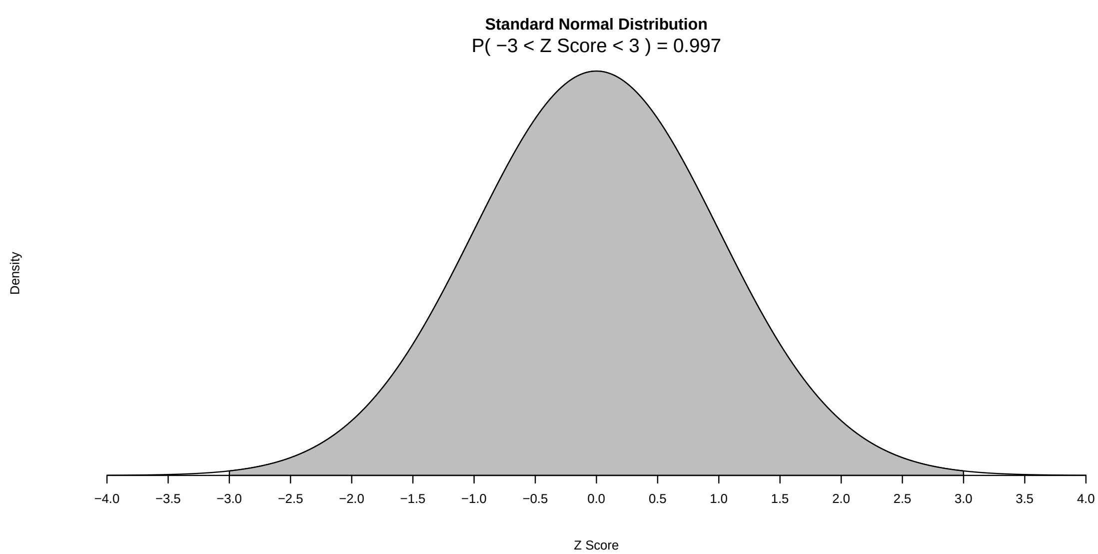
Characteristics of Normal Probability Distribution

1. The distribution is **symmetric**; its skewness measure is zero.
2. The entire family of normal probability distributions is defined by its mean μ and its standard deviation σ.
3. The highest point on the normal curve is at the mean, which is also the median and mode.
4. The mean can be any numerical value: negative, zero, or positive.
5. The standard deviation determines the width of the curve: larger values result in wider, flatter curves.
6. Probabilities for the normal random variable are given by areas under the curve. The total area under the curve is 1 (0.5 to the left of the mean and 0.5 to the right).
7. A random variable having a normal distribution with a mean of 0 and a standard deviation of 1 is said to have a **standard normal probability distribution**.

Empirical Rule of Normal Distribution:


Standard Normal Distribution
P(−1 < Z Score < 1) = 0.683
Density
−4.0 −3.5 −3.0 −2.5 −2.0 −1.5 −1.0 −0.5 0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0
Z Score


Standard Normal Distribution
P(−2 < Z Score < 2) = 0.954
Density
−4.0 −3.5 −3.0 −2.5 −2.0 −1.5 −1.0 −0.5 0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0
Z Score


Standard Normal Distribution
P(−3 < Z Score < 3) = 0.997
Density
−4.0 −3.5 −3.0 −2.5 −2.0 −1.5 −1.0 −0.5 0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0
Z Score

Calculate the Z Score

A Z-score is a numerical measurement that describes a value's relationship to the mean of a group of values. Z-score is measured in terms of standard deviations from the mean. If a Z-score is 0, it indicates that the data point's score is identical to the mean score. A Z-score of 1.0 would indicate a value that is one standard deviation from the mean. Z-scores may be positive or negative, with a positive value indicating the score is above the mean and a negative score indicating it is below the mean.

We can think of z as a measure of the number of standard deviations x is from μ. Using the following formula, we can convert a numerical measurement to the standardized z score:

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

Example:

Pep Zone sells auto parts and supplies including a popular multi-grade motor oil. When the stock of this oil drops to 20 gallons, a replenishment order is placed. The store manager is concerned that sales are being lost due to stockouts while waiting for a replenishment order. It has been determined that demand during replenishment lead-time is normally distributed with a mean of 15 gallons and a standard deviation of 6 gallons. The manager would like to know the probability of a stockout during replenishment lead-time. In other words, what is the probability that demand during lead-time will exceed 20 gallons?

$$P(x > 20) =?$$

Step 1: Convert x to the standard normal distribution

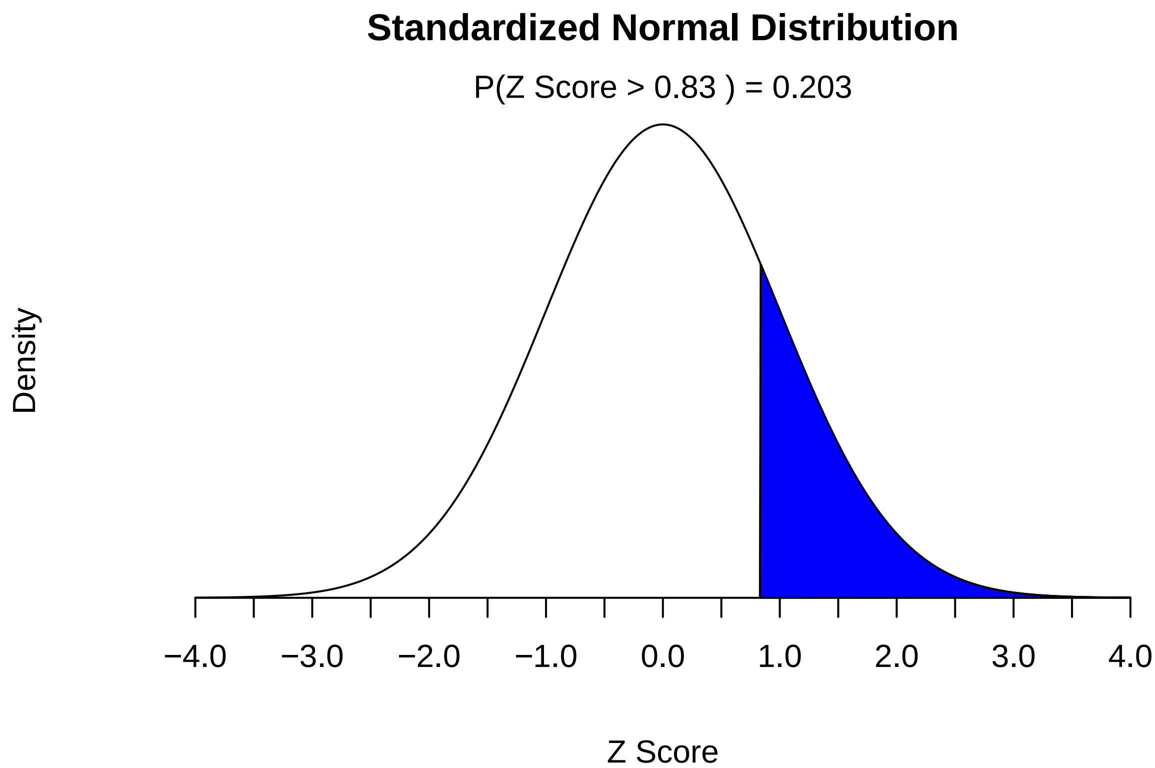
$$z = \frac{20 - 15}{6} = 0.83$$

Step 2: Find the area under the standard normal curve to the right of z = 0.83

Use the Probability Table for the Standard Normal Distribution

$$P(z > 0.83) = 0.2033 \textit{ or } 20.33\%$$

Graphical Solution:



If the manager of Pep Zone wants the probability of a stockout during replenishment lead-time to be no more than .05, what should the reorder point be? (Hint: Given a probability, we can use the standard normal table in an inverse fashion to find the corresponding z value.)

$$x = \mu + z_{0.05}\sigma$$

$$x = 15 + 1.645 * 6$$

$$x = 24.87 \; or \; 25$$

A reorder point of 25 gallons will place the probability of a stockout during lead time at (slightly less than) 0.05 or 5%.

Exponential Probability Distribution

The exponential probability distribution is useful in describing the time it takes to complete a task. The exponential random variables can be used to describe:

- Time between vehicle arrivals at a toll booth
- Time required to complete a questionnaire
- Distance between major defects in a highway

In waiting line applications, the exponential distribution is often used for service time.

A property of the exponential distribution is that the mean and standard deviation are equal. The exponential distribution is skewed to the right. Its skewness measure is 2.

Exponential Probability Density Function:

$$f(x) = \frac{1}{\mu} e^{-x/\mu} \ for \ x \geq 0$$

where:

μ = expected value or mean

e = 2.71828...

Relationship Between the Poisson and Exponential Distribution:

The Poisson and Exponential distributions are closely related to each to each other. The Poisson distribution provides an appropriate description of the number of occurrences per interval. While the exponential distribution provides an appropriate description of the length of the interval between occurrences. Just so, the Poisson distribution deals with the number of occurrences in a fixed period, and the exponential distribution deals with the time between occurrences of successive events as time flows by continuously.

Just so, the Poisson distribution deals with the number of occurrences in a fixed period, and the exponential distribution deals with the time between occurrences of successive events as time flows by continuously.