CEE 631
CONSTRUCTION DECISIONS UNDER UNCERTAINTY

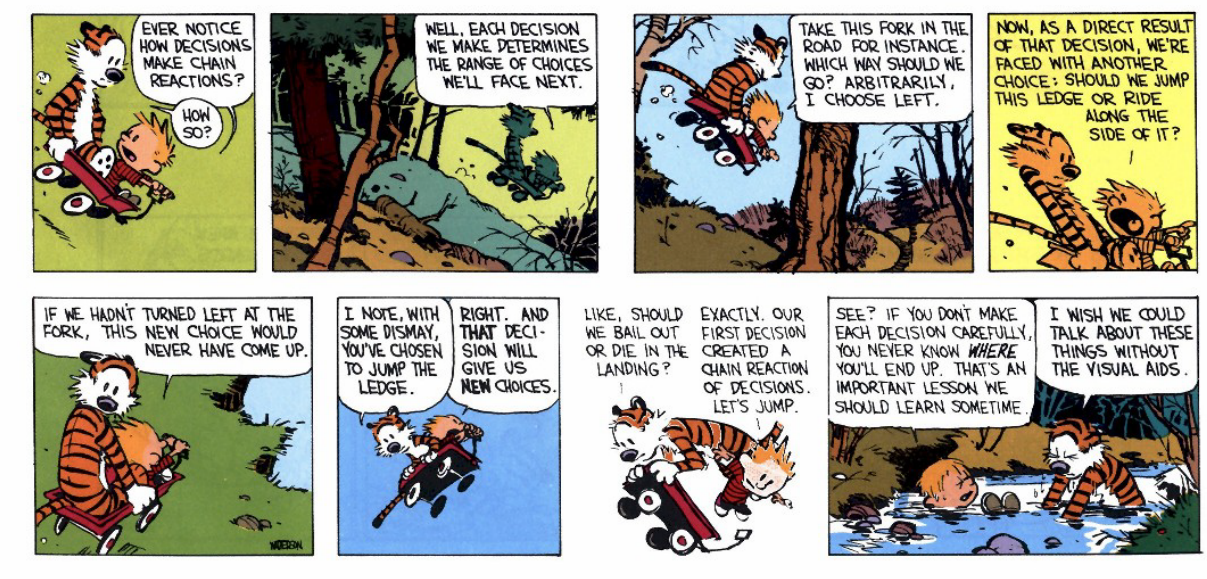
CLASS NOTES
PART 1

Review of Probability & Statistics



© 2024 Photios G. Ioannou, Ph.D., Dipl.Eng., F.ASCE
Professor of Civil and Environmental Engineering
Construction Engineering & Management
University of Michigan
Ann Arbor, Michigan 48109-2125

Decision Making Explained



Decisions Must be Made – One Way or Another



Elements of Decision Making

A decision is a choice from among two or more alternative actions. The most attractive alternative is the one which has the highest value for the decision maker. In general, two different decision makers, faced with a seemingly identical situation, may assign different values to the outcomes of the decision alternatives for a variety of reasons. One of those reasons is one's personal wealth. Another is aversion to risk. Another is each decision maker's state of information. Another is the degree of control that each decision maker can exercise in determining outcomes. And both information and control as well as prior experience can have a profound influence on the assessment of probabilities for the possible outcomes. So, each decision, when viewed from the point of view of the decision maker, is essentially unique.

At the beginning of this course, we will adopt the most common decision criterion for making rational decisions, i.e., the maximization of expected value. As we will see later, this approach can be extended, in a straightforward manner, to cases where the decision maker dislikes risk.

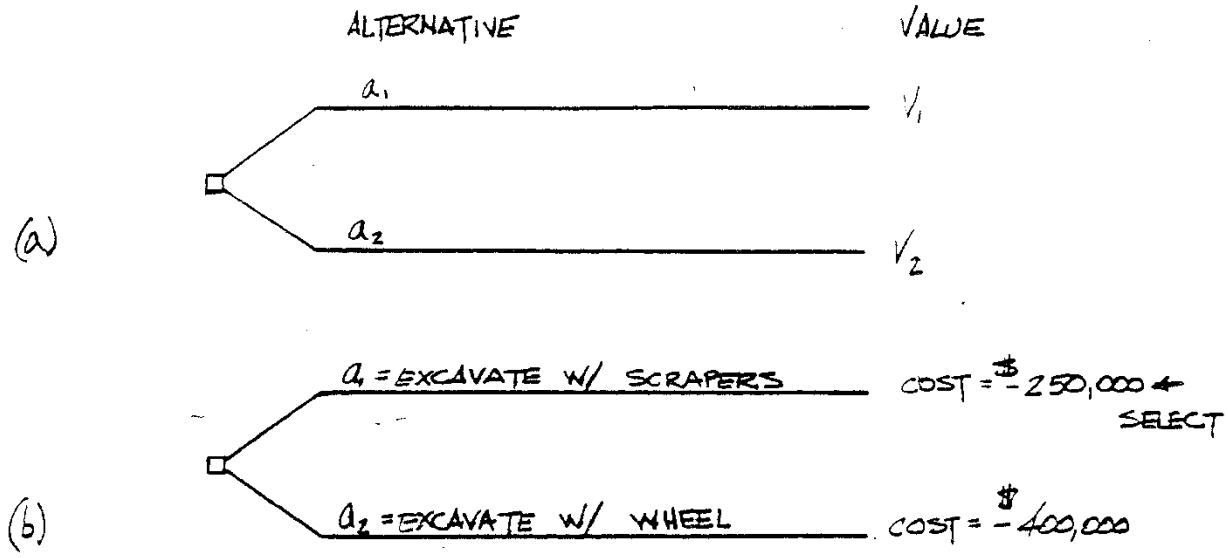
To determine the alternative which has *the highest expected value*, the decision maker must first determine the expected value *of each alternative*. This calculation requires the following five elements:

1. A list of the possible alternatives.
2. A list of the possible outcomes of each alternative.
3. Probability of occurrence of each outcome.
4. Value to the decision maker of each outcome.
5. Ability to combine these elements to determine expected value.

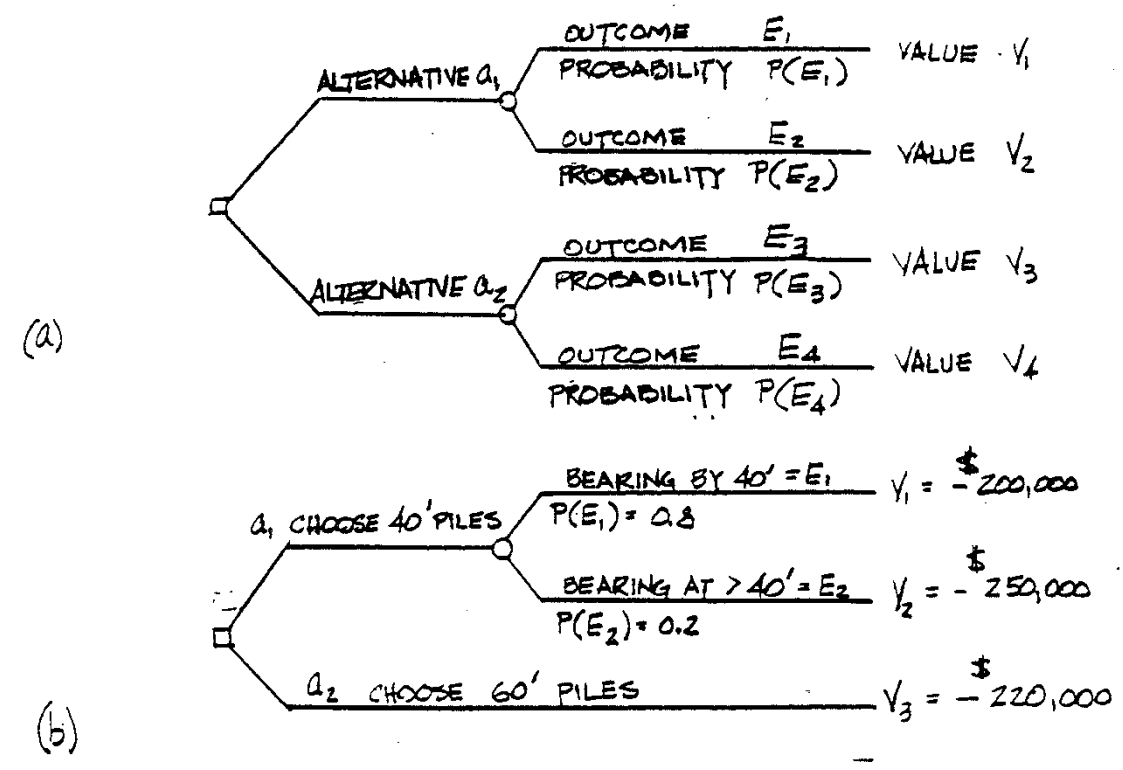
In construction, a decision maker often has considerable knowledge (from past experience and from studying the projects available) concerning the first four elements listed above. One of our objectives in this course is to develop the ability to combine these elements so that the decision maker can use his knowledge about a decision situation to make good decisions. This ability in no way replaces the need for that knowledge, but instead, makes better use of it. It is often claimed that the most important requirement of a construction engineer is the ability to make timely decisions. In this course, we will add the requirement that decisions should be made well.

The simplest form of decision assumes that there is a single outcome to each alternative. The decision is then reduced to the selection of the best outcome. There is a strong temptation to reduce most decisions to this form because it is simple and because the decision maker lacks time or ability to be more accurate. Another form of this problem is the temptation to select an alternative with a known attractive outcome, without considering alternatives with more attractive but uncertain outcomes. In fact, because decisions always involve future actions, all outcomes are uncertain, and a decision maker should take care not to oversimplify them if their outcomes are at all important.

Because the outcomes are uncertain, a decision can be viewed as a choice from among alternative "experiments". Put simply, a person chooses an "experiment" from among several alternative experiments and awaits its uncertain outcome.



Simple Decision Alternatives with Certain Outcomes



Simple Decision Alternatives with Uncertain Outcomes

Sample Space, Events

The collection of all possible outcomes of an experiment (such as a decision alternative) is its *sample space,* S, which is made of *sample points* (or *simple events*), E_i. Thus, the sample space, $S=\{E_i\}$, is the set of all sample points, E_i. Performing the experiment is often called *sampling the experiment* and the outcome of one sampling is a *sample*. Clearly, a sample must be one of the simple events E_i that make up the sample space S.

Suppose that on Wednesday afternoon, a superintendent contacts the business agent of the carpenters' union and asks for five carpenters to come for work on Friday morning. The superintendent feels that the number of carpenters who actually show up for work on Friday morning has a sample space containing the numbers 0, 1, 2, 3, 4, and 5 as follows:

$$S = \{0,1,2,3,4,5\} = \{E_0, E_1, E_2, E_3, E_4, E_5\}$$

The simplest possible outcomes of an experiment are called *simple events* and define the individual sample points in the sample space S. The example above has six such *simple events*: E_0, E_1, E_2, E_3, E_4, E_5.

A *compound event*, A, is a set of *several* simple events E_i from within a sample space S, such as $A=\{4,5\}=\{E_4,E_5\}$. A compound event *A occurs* whenever *one* of its simple events E_i occurs.

Many compound events can be defined for the same sample space. For example, compound event A could be that "4 or 5 carpenters arrive". Another compound event B would be that "fewer than 3 arrive". A third event C would be that "an odd number showed up". Event D is that "at most one carpenter shows up". Each of these compound events is shown below.

$$A = \{4,5\} = \{E_4, E_5\}$$
$$B = \{0,1,2\} = \{E_0, E_1, E_2\}$$
$$C = \{1,3,5\} = \{E_1, E_3, E_5\}$$
$$D = \{0,1\} = \{E_0, E_1\}$$

Since compound events are *sets*, the algebra of compound events uses the basic *set operators*:

Intersection:	$A \cap C = \{5\} = \{E_5\}$
Union:	$A \cup C = \{1,3,4,5\} = \{E_1, E_3, E_4, E_5\}$
Complement:	$A^C = A' = \{0,1,2,3\} = \{E_0, E_1, E_2, E_3\}$

When there is no access to special symbols (e.g., ∩), *algebraic* operators may be used as *set* operators (however, when possible, this practice should be avoided to avoid confusion):

Intersection:	$AB = A \cap B$
Union:	$A + B = A \cup B$
Difference:	$A - B = A(-B) = AB' = A \cap B' = A \cap B^C$
Negation	$-A = A' = A^C$

The *intersection* (or *intercept*) of events A and B is the occurrence of a sample point that belongs to both A and B. The *union* of events A and B is the occurrence of an event that belongs to at least one of them. The *complement* of an event A is its non-occurrence.

Events A and B are called *mutually exclusive* (ME), or *disjoint*, if they cannot both occur at the same time, i.e., if they have no sample points in common (their intersection is the null set):

$$A \cap B = \varnothing$$

Events A, B,…, are called *collectively exhaustive* (CE) if their union is the sample space S. For example, the events A, B, C defined above are collectively exhaustive because:

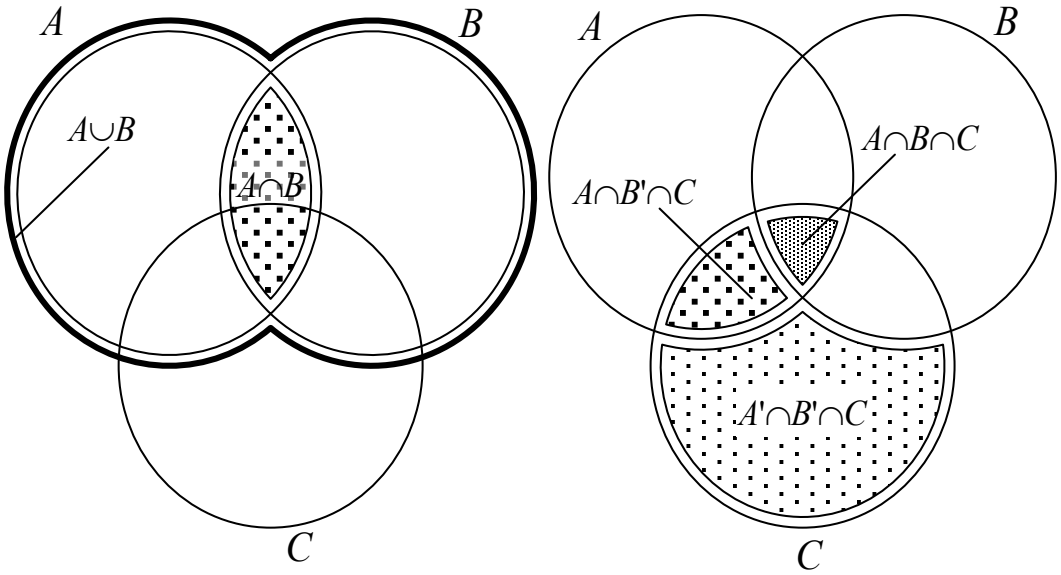
$$A \cup B \cup C = S$$

If the intersection of two compound events is equivalent to one of these events, that event is said to be contained in the other. For example, $B \cap D = \{E_0, E_1\} = D$. Therefore:

B is a superset of D (B contains D):	$B \supset D$
D is a subset of B (D is contained in B)	$D \subset B$

Sometimes the definition of an event can be complex, such as "an odd number less than 4 arrive": $C \cap A' = \{E_1, E_3\}$. In all cases, the description of such an event must be explicit and accurate (not approximate). A complex event may have several possible correct definitions. Also, there may be several other definitions that may appear to be correct but are in fact in error. Great care should therefore be taken in defining events accurately. Otherwise, if definitions are incorrect, all calculations which follow are meaningless. This is a common source of errors.

A great aid is to show the combinations of events graphically using *Venn* diagrams. Each compound event A, B, C, is shown by a circle which is overlapped with each of the other events, as necessary. The overlapped areas represent the sample points that belong to more than one of the compound events. This representation is very helpful for understanding event algebra.



PERMUTATIONS, VARIATIONS, COMBINATIONS

(1) Permutations

(a) A permutation is an arrangement of n elements. The number of all possible permutations of n different elements is

$$_nP_n = n(n-1)(n-2)\cdots(3)(2)(1) = n!$$

(b) The number of all different permutations of n elements, among which there are a elements of equal value, is

$$_aP_n = \frac{n(n-1)(n-2)\cdots(3)(2)(1)}{a(a-1)(a-2)\cdots(3)(2)(1)} = \frac{n!}{a!}$$

(c) The number of all different permutations of n elements, among which there are a elements of one equal value and b elements of another equal value, is

$$_{a,b}P_n = \frac{n(n-1)(n-2)\cdots(3)(2)(1)}{a(a-1)(a-2)\cdots(3)(2)(1)b(b-1)(b-2)\cdots(3)(2)(1)} = \frac{n!}{a!b!}$$

(2) Variations

A variation is an arrangement of n elements into a sequence of k terms. The number of all possible variations is

$$_kV_n = \frac{n(n-1)(n-2)\cdots(3)(2)(1)}{(n-k)(n-k-1)(n-k-2)\cdots(3)(2)(1)} = \frac{n!}{(n-k)!} = \binom{n}{k}k!$$

(3) Combinations

A combination is an arrangement (without repetition) of n elements into a sequence of k terms. The number of all possible combinations is

$$_kC_n = \frac{n(n-1)(n-2)\cdots(n-k+2)(n-k+1)}{(n-k)(n-k-1)(n-k-2)\cdots(3)(2)(1)} = \frac{n!}{(n-k)!k!} = \binom{n}{k}$$

(4) Table – Example

Permutations	Elements A, B, C	$n = 3$
	$ABC \quad BCA \quad CAB$ $ACB \quad BAC \quad CBA$	$P_3 = (3)(2)(1) = 6$
Permutations	**Elements A, A, C**	**$n = 3 \quad a = 2$**
	$AAC \quad ACA \quad CAA$	$_2P_3 = \frac{(3)(2)(1)}{(2)(1)} = 3$
Variations	**Elements A, B, C**	**$n = 3 \quad k = 2$**
	$AB \quad BC \quad CA$ $BA \quad CB \quad AC$	$_2V_3 = \frac{(3)(2)(1)}{1} = 6$
Combinations	**Elements A, B, C**	**$n = 3 \quad k = 2$**
	$AB \quad BC \quad CA$	$_2C_3 = \frac{(3)(2)(1)}{(2)(1)} = 3$

© 2024 *Photios G. Ioannou*

Probability Theory

Associated with the outcome of a random experiment is a measure of its likelihood of occurrence which is its probability of occurrence. The flipping of a coin will result in an outcome of heads with a probability of 1/2 and of tails with a probability of 1/2. In a discrete random variable, which would have a discrete sample space, each sample point would have a probability of occurrence.

The modern axiomatic foundation of probability theory requires only three axioms as follows:

1. The probability of any event A is defined as $P[A]$, such that,

$$0 \le P[A] \le 1 \tag{1}$$

2. The probability of the certain event S containing all sample points is 1,

$$P[S] = 1 \tag{2}$$

3. The probability of an event which is the union of two mutually exclusive events is the sum of their probabilities,

$$P[A \cup B] = P[A] + P[B] \tag{3a}$$

Because all sample points E_i are mutually exclusive and collectively exhaustive, the last axiom can alternatively be written as follows,

$$\sum_{\text{all } i} P[E_i] = 1 \tag{3b}$$

Based on these axioms it is possible to derive several simple propositions, such as the probability of the union of two events,

$$P[A \cup B] = P[A] + P[B] - P[A \cap B] \tag{4}$$

Calculation of Probability of an Event by Brute Force

1. Describe the event, A, for which the probability is desired.

2. Describe the simple events, i.e., the mutually exclusive sample points, E_i, which make up the event A. These may be compound events made up of joint occurrences of simple events. Unless the event A is certain with a probability one, the sample points will only be a part of some sample space, of which the remainder is A'. As an alternative to describing the sample points which make up A, one can choose to describe the sample points which make up A', remembering that $P[A] = 1 - P[A']$.

3. Determine the probability $P[E_i]$ of each sample point in A (or A'). This requires the input of probabilities of simple events and conditional probabilities of compound events.

4. Calculate the probability of event A by summing the probabilities of the mutually exclusive sample points.

$$P[A] = \sum_{all\ i} P[E_i]$$

Example Calculation of Probabilities of Events

The relationships of basic probability theory are used for the different combinations of events. The sample points can be considered events. Since the occurrence of one sample point for an experiment means that none of the others occur (for that experiment), each is mutually exclusive of the others. The most straightforward (and foolproof) method of calculating the probability of an event is to follow the equation for disjoint events and to sum the probabilities of the sample points which make it up.

$$P[A] = \sum_{\text{all } E_i \text{ in } A} P[E_i]$$

Example: A superintendent requests of the union business agent that 5 carpenters be assigned for a Friday morning. For the random variable of number of carpenters which are assigned, the following table of probabilities is estimated.

Number assigned, E_i	E_0	E_1	E_2	E_3	E_4	E_5
$P[E_i]$	0.15	0.05	0.05	0.15	0.2	0.4

With $A = \{E_4, E_5\}$ $\quad P[A] = P[E_4] + P[E_5] = 0.2 + 0.4 = 0.6$

With $C = \{E_1, E_3, E_5\}$; $\quad P[C] = P[E_1] + P[E_3] + P[E_5] = 0.05 + 0.15 + 0.4 = 0.6$

In order to use this "brute force" technique on an event which is the combination of two events, one must use their common sample points. If the two events are drawn from the same sample space, the probability of the occurrence of their combination is the sum of the probabilities of the sample points within that combination.

Example. For the probabilities given above for the possible assignment of carpenters and for $A = \{E_4, E_5\}$, and $C = \{E_1, E_3, E_5\}$,

$$A \cap C = \{E_5\} \quad \Rightarrow \quad P[A \cap C] = P[E_5] = 0.4$$

Example. For $A = \{E_4, E_5\}$, we have $A' = \{E_0, E_1, E_2, E_3\}$:

$$A' \cap C = \{E_1, E_3\} \quad \Rightarrow \quad P[A' \cap C] = P[E_1] + P[E_3] = 0.05 + 0.15 = 0.2$$

Example. The probability of $A \cup C = \{E_1, E_3, E_4, E_5\}$ can be computed now in two ways,

$$P[A \cup C] = P[E_1] + P[E_3] + P[E_4] + P[E_5] = 0.05 + 0.15 + 0.2 + 0.4 = 0.8$$

$$P[A \cup C] = P[A] + P[C] - P[A \cap C] = 0.6 + 0.6 - 0.4 = 0.8$$

If the combination is of events from different sample spaces, then this summation must be from their *joint* sample space for their occurrence together. This requires the calculation of the probabilities of the joint occurrence of the sample points, a straightforward operation if the sample points are independent.

© 2024 *Photios G. Ioannou*

Conditional probability

The probability that event A will occur if event B is known to occur is the probability of A given B, $P[A \mid B]$, and defined as,

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} \tag{5}$$

The sample space of event $A \mid B$ is reduced from the entire sample space, S, for event A, to the conditional sample space, B. Thus, the conditional line "|" reduces all space S for event A to that subset which is also contained in B.

The definition of conditional probability provides another way to calculate the probability for the intersection of two or more events. From the equation above, it is evident that,

$$P[A \cap B] = P[A \mid B]P[B] = P[B \mid A]P[A] \tag{6}$$

Similarly, the intersection of three events A, B and C can be computed by considering either $A \cap B$, or $B \cap C$ as a single event and using the formula for two events twice,

$$\begin{aligned} P[A \cap B \cap C] &= P[A \mid (B \cap C)]P[B \cap C] = P[A \mid B \cap C]P[B \mid C]P[C] \\ &= P[(A \cap B) \mid C]P[C] = P[A \mid B \cap C]P[B \mid C]P[C] \end{aligned}$$

This result can be generalized to calculate the probability of the intersection of any number of events as follows,

$$\begin{aligned} P[A \cap B \cap C \cap \cdots \cap Y \cap Z] = P[&A \mid B \cap C \cap \cdots \cap Y \cap Z] \\ &\cdot P[B \mid C \cap \cdots \cap Y \cap Z] \\ &\cdot P[Y \mid Z] \\ &\cdot P[Z] \end{aligned}$$

Independent Events

If event A and event B are independent, then knowledge of the occurrence (or nonoccurrence) of one of them has no influence on the probability of occurrence of the other.

Mathematically, events A and B are independent if

$$P[A \mid B] = P[A] \quad \text{or} \quad P[B \mid A] = P[B] \tag{7}$$

From here, if A and B are independent it follows that,

$$P[A \cap B] = P[A]P[B] \tag{8}$$

Ask Marilyn®

BY MARILYN VOS SAVANT



A shopkeeper says she has two new baby beagles to show you, but she doesn't know whether they're male, female or a pair. You tell her that you want only a male, and she telephones the fellow who's giving them a bath. "Is at least one a male?" she asks. "Yes," she informs you with a smile. What is the probability that the *other* one is a male?

—Stephen I. Geller, Pasadena, Calif.

The key to many probability problems is accuracy and precision in the definition of the events of interest. In this example, the events of interest are A, and B|A:

P[A] = P[At least one of the two puppies is a male]
= 1 - P[Both puppies are female] = 3/4.

P[B|A] = P[The *other* puppy is a male | A]
= P[The *other* puppy is a male and A] / P[A]
= P[Both puppies are male] | P[A]
= (1/4)/(3/4)=1/3

Notice that the word "*other*" can only define a *conditional* event B|A. Otherwise, it is meaningless.

Ask Marilyn®

BY MARILYN VOS SAVANT



This is regarding the problem where a shopkeeper has two baby beagles, but she doesn't know if they're male, female or a pair. You come into the shop to buy a pup, but you want only a male, so she phones the fellow who is giving the beagles a bath and asks, "Is at least one a male?" After he gives her the answer, she tells you, "Yes!"

What's the probability that the other one is a male?

You answered, "One out of three." There are only three possible explanations for this: 1) You are considering the possibility that the second puppy has been neutered, but because you did not make this clear, you leave open the next explanation. 2) You are (still) confused about probability. If so, I suggest you brush up on this information. 3) Your incorrect responses are intentional, done as a means to solicit mail. If so, there has to be a better way.

—James Larsen, Ph.D., Wright State University

I disagree with your answer. The observer knows the first dog is a male; the probability that the other is a male is 50-50.

—Richard Jones, Ph.D., University of New Haven

I believe you're wrong! Before you know that one of the baby beagles is a male, the chances are one out of three. But once the above fact is known, the chances change to 50-50.

—Edward Weiss, Ph.D., Bethel College

Okay. But here's what is puzzling to me: If there were only one beagle pup, the probability of it being male would be one out of two. What can explain why the presence of another beagle pup should affect the probability that the pup in question is a male? I know this is cute, but how far away does the second beagle have to be in order for it not to affect the sex of the first beagle?

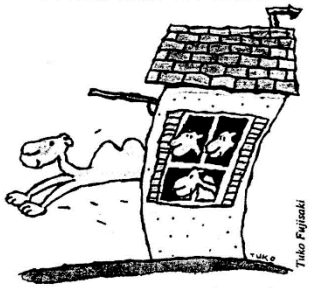
—Steve Marx, Worcester, Mass.

The original answer is correct. We didn't define a "first beagle" or a "second beagle," and so either beagle can be in a doghouse on the moon, and it still would affect the outcome.

If we could shake a pair of puppies out of a cup the way we do dice, there are four ways they could land: male/female or female/male or male/male or female/female. So there are three ways in which at least one of them could be a male. And since the partner of a male in those three is either a female, a female or another male, the chances of that partner being a male are one out of three.

The key is that we didn't specify which beagle was a male, so it can be either one. If we'd said instead, "The one nibbling your ankle is a male—what are the chances that the one sleeping is a male?" the chances would be 50-50. But we just specified that "at least one" was a male, so we don't know which it is and whether its partner is awake or asleep. And this means the chances of that partner being a male are only one out of three.

Here's A Brainteaser From Me To You:



Suppose there are 11 camels in your garage. Six of them have two humps, and five of them have one hump. You've asked your best friend to go for a ride, but the animals should match. You don't want to look silly! However, as camels are pretty hard to catch, you're just going to open the side door and stand out of the way in back of it. Assuming that they'll shoot out the door at the rate of one per second, how long is the minimum you should wait before you slam the door shut to make sure you find a pair of matching camels in the backyard? (*Answer will appear in next week's column.*)

Example: Baby-Beagles–Marilyn Vos Savant

Unconditional sample space:

{(M,M), (M,F), (F,M), (F,F)}

Conditional sample space given that at least one puppy is a male:

{(M,M), (M,F), (F,M)}

Of the three simple events, one simple event has two males:

P[Other one is male | at least one male] = 1/3

So, Marilyn is correct. But the problem was stated in a tricky way.

Notice that:

P[Other one is male | first one is a male] = 1/2

P[Other one is male | this one is a male] = 1/2

In the problem, as stated, we do not know which of the two puppies is a male. All we know is that there is at least one. This information is not as strong as knowing which of the two is indeed a male. Thus, the probability of the other being a male is only 1/3, instead of 1/2.

You can verify that the correct answer is 1/3 by tossing two coins and observing the outcomes. Assume that "heads" is the same as "male."

Example: Camels in the Garage–Marilyn Vos Savant

There are 11 camels, 6 with two humps, and 5 with one hump. They exit the garage one every 1 second in random order. You cannot see the exiting camels.

(a) What is the minimum you have to wait before closing the door to be sure you get a matching pair of camels? The answer is 3 seconds. This problem has nothing to do with probability!

(b) However, if you can see the camels exiting through the door, then the minimum time you have to wait, X, is either 2 or 3 seconds. Here is how to find the expected value, E[X].

Sample space for X is {2, 3}

$$\begin{aligned} P[X=2] &= P[(1,1)\cup(2,2)] = P[(1,1)] + P[(2,2)] \\ &= \tfrac{5}{11}\times\tfrac{4}{10} + \tfrac{6}{11}\times\tfrac{5}{10} = \tfrac{2}{11} + \tfrac{3}{11} = \tfrac{5}{11} \\ P[X=3] &= P[(1,2)\cup(2,1)] = P[(1,2)] + P[(2,1)] \\ &= \tfrac{5}{11}\times\tfrac{6}{10} + \tfrac{6}{11}\times\tfrac{5}{10} = \tfrac{3}{11} + \tfrac{3}{11} = \tfrac{6}{11} \\ E[X] &= 2\times\tfrac{5}{11} + 3\times\tfrac{6}{11} = \tfrac{28}{11} = 2.54\,\text{sec} \end{aligned}$$

Conditional Probability Example

Let D be the demand level for water and S represent the supply level for water and assume that the following probabilities are available:

$P[D,S]$	S	
D	50	60
30	0.2	0.1
40	0.3	0.05
50	0.1	0.25

$$P[D=30]=P[30,50]+P[30,60]=0.2+0.1=0.3$$

$$P[S=50]=P[30,50]+P[40,50]+P[50,50]=0.2+0.3+0.1=0.6$$

$$P[S=50\,|\,D=30]=\frac{P[30,50]}{P[D=30]}=\frac{0.2}{0.3}=\frac{2}{3}$$

Notice that events $(S=50)$ and $(D=30)$ are not independent because,

$$P[S=50\,|\,D=30]=\frac{2}{3}\neq P[S=50]=0.6$$

Similarly, the conditional demand probabilities can be computed as follows,

$$P[D=30\,|\,S=50]=\frac{P[30,50]}{P[S=50]}=\frac{0.2}{0.6}=\frac{1}{3} \qquad P[D=30\,|\,S=60]=\frac{P[30,60]}{P[S=60]}=\frac{0.1}{0.4}=\frac{1}{4}$$

$$P[D=40\,|\,S=50]=\frac{P[40,50]}{P[S=50]}=\frac{0.3}{0.6}=\frac{1}{2} \qquad P[D=40\,|\,S=60]=\frac{P[40,60]}{P[S=60]}=\frac{0.05}{0.4}=\frac{1}{8}$$

$$P[D=50\,|\,S=50]=\frac{P[50,50]}{P[S=50]}=\frac{0.1}{0.6}=\frac{1}{6} \qquad P[D=50\,|\,S=60]=\frac{P[50,60]}{P[S=60]}=\frac{0.25}{0.4}=\frac{5}{8}$$

$P[D\,\|\,S]$	S	
D	50	60
30	1/3	1/4
40	1/2	1/8
50	1/6	5/8

Note: The notation $P[D=d\cap S=s]$ is both verbose and cumbersome and for this reason the example above uses the convenient shorthand notation $P[D,S]=P[D=d\cap S=s]$. Shortly, we will introduce the notation for random variables that combines precision with convenience: $p_{D,S}[d,s]=P[D=d\cap S=s]$

Pairwise & Mutual Independence

A set of events $\{A_1, A_2, \ldots, A_n\}$ is said to be *pairwise independent* if for any pair A_i and A_j

$$P[A_i \cap A_j] = P[A_i]P[A_j]$$

A set of events $\{A_1, A_2, \ldots, A_n\}$ is said to be *mutually independent* if for *any* subset $\{A_i, A_j, \ldots, A_m\}$ of these events we have

$$P[A_i \cap A_j \cap \ldots \cap A_m] = P[A_i]P[A_j]\ldots P[A_m]$$

Theorems

If a set of events are mutually independent, they must also be pairwise independent.

If a set of events are pairwise independent, they do not have to be mutually independent.

The single fact that $P[A_1 \cap A_2 \cap \ldots \cap A_n] = P[A_1]P[A_2]\ldots P[A_n]$ *does not necessarily imply that the events* $\{A_1, A_2, \ldots, A_n\}$ *are mutually independent.*

Example

CW1: Cement weight measured correctly
CW2: Cement weight measured 20 lbs too low
AW1: Aggregate weight measured correctly
AW2: Aggregate weight measured 20 lbs too high

	Aggregate Weight	
Cement Weight	AW1	AW2
CW1	$E_{1,1}$	$E_{1,2}$
CW2	$E_{2,1}$	$E_{2,2}$

Define the following compound events,

$A = \{E_{1,1},\ E_{1,2}\}$ = No error in measurement of cement weight

$B = \{E_{1,1},\ E_{2,1}\}$ = No error in measurement of aggregate weight

$C = \{E_{1,2},\ E_{2,1}\}$ = Only one of cement or aggregate is measured incorrectly

Assuming that all events $E_{1,1}$, $E_{1,2}$, $E_{2,1}$, $E_{2,2}$, have equal probability = 0.25

$$P[A] = P[B] = P[C] = 0.25 + 0.25 = 0.5$$

$P[A \cap B] = P[E_{1,1}] = 0.25 = P[A]P[B] \Rightarrow$ Events A and B are pairwise independent

$P[B \cap C] = P[E_{2,1}] = 0.25 = P[B]P[C] \Rightarrow$ Events B and C are pairwise independent

$P[C \cap A] = P[E_{1,2}] = 0.25 = P[C]P[A] \Rightarrow$ Events C and A are pairwise independent

However,

$$P[A \cap B \cap C] = P[\varnothing] = 0 \neq P[A]P[B]P[C] = 0.125$$

Thus, even though events A, B, C are pairwise independent, they are not mutually independent.

Language of Probabilities

Translation is required between words and the expressions of events and their probabilities. Some key words indicate the appropriate translation.

The event that A occurs given that B occurs (has occurred) (will occur) is a conditional event and is written as $A \mid B$. The conditional event $A \mid B$ can be expressed by a variety of statements, such as,

		(given that)		
	(occurs)	(if)		(occurs)
A	(has occurred)	(if it is known that)	B	(has occurred)
	(will occur)	(knowing that)		(will occur)
		(when)		

In probability problems, the event that A *or* B occurs is the union event $A \cup B$. This event $A \cup B$ can also be expressed as A occurs *and/or* B occurs; or, o*ne or more* of A and B occurs. The event where *exactly one of* A and B occurs (but not both) is called an *exclusive or* event and must be written as $(A \cup B) \cap (A \cap B)'$. However, it does not appear often in probability problems. If A and B are mutually exclusive, then the *exclusive or* statement and the more common *or* statement are the same and mean $\mathrm{A} \cup \mathrm{B}$.

The event that A *and* B occur is $A \cap B = AB$ which is the event that *both* A *and* B occur.

Examples.

- One worker is sick (S_1), *and* two are injured (I_2): $S_1 \cap I_2$, or $S_1 I_2$.
- One worker is sick (S_1), *and/or* two are injured (I_2): $S_1 \cup I_2$.
- Either one worker is sick (S_1), *or* two are injured (I_2) (*but not both*): $(S_1 \cup I_2)(S_1 I_2)'$.
- One worker is sick (S_1), *if* two are injured (I_2): $S_1 \mid I_2$.
- One worker is sick (S_1), *and* two are injured (I_2), *when* ten were assigned (A_{10}): $S_1 I_2 \mid A_{10}$.
- One worker is sick (S_1), and two (I_2) or three (I_3) are injured, *when* nine (A_9) or ten (A_{10}) were assigned: $S_1(I_2 \cup I_3) \mid (A_9 \cup A_{10})$.
- One worker is sick (S_1) and two are injured (I_2) or three are injured (I_3) is ambiguous because it could be either $S_1(I_2 \cup I_3)$ or it could mean $(S_1 I_2) \cup I_3$. These are not the same, since by de Morgan's formulas, $S_1(I_2 \cup I_3) = (S_1 I_2) \cup (S_1 I_3) \neq (S_1 I_2) \cup I_3$ because $S_1 I_3 \neq I_3$.

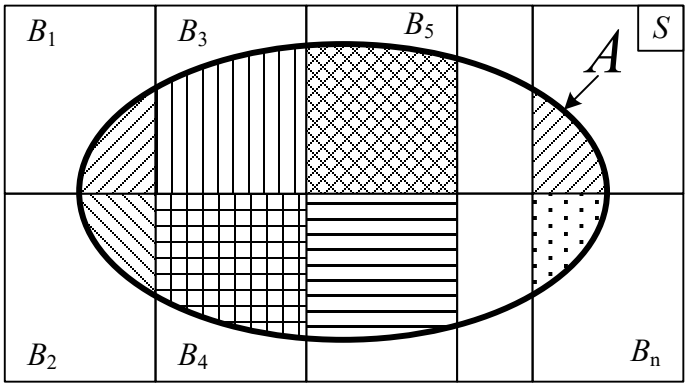
Total Probability Theorem

Let A be the event of interest and $B_1, B_2, \ldots, B_n$ be a collection of mutually exclusive and collectively exhaustive (MECE) events:

$$B_i \cap B_j = \varnothing \quad (i \neq j)$$

$$\bigcup_{i=1}^{n} B_i = S$$

The Venn diagram below shows the sample space S partitioned into events $B_1, B_2, \ldots, B_n$ and their intersection with A.



Total Probability Theorem: The probability of event A is given by the sum of the shaded areas:

$$\begin{aligned} P[A] &= P[A \cap B_1] + P[A \cap B_2] + \ldots + P[A \cap B_n] \\ &= \sum_{all\ i} P[A \cap B_i] = \sum_{all\ i} P[A \mid B_i] P[B_i] \end{aligned}$$

Corollary: The Total Probability Theorem provides an alternative way to calculate $P[A \cap B]$.

$$\begin{aligned} P[A \cap B] &= P[A \cap (B_1 \cup B_2 \cup \ldots \cup B_n)] = P[A \cap B_1] + P[A \cap B_2] + \ldots + P[A \cap B_n] \\ &= \sum_{all\ i} P[A \cap B_i] = \sum_{all\ i} P[A \mid B_i] P[B_i] \end{aligned}$$

Here, events $B_1, B_2, \ldots, B_n$ are mutually exclusive and their union is event B, a subset of S.

$$B_i \cap B_j = \varnothing \quad (i \neq j)$$

$$\bigcup_{i=1}^{n} B_i = B \subseteq S$$

Example. A soils engineer reports that a stratum of dense clay in the immediate area varies in depth below the surface from about 15' to 25' with probabilities:

$$P[D=15']=0.2,\ P[D=20']=0.5,\ P[D=25']=0.3$$

He also estimates that driven piles will reach the required resistance with the following probabilities.

$P[Y \mid D]$	Depth to resistance = Y (ft)			
Clay Depth, D (ft)	$y_1=20$	$y_2=25$	$y_3=30$	$y_4=35$
$d_1=15$	0.2	0.6	0.2	0
$d_2=20$	0	0.3	0.6	0.1
$d_3=25$	0	0.1	0.7	0.2

The probability of needing to drive exactly 30' to resistance can be calculated using the total probability theorem:

$$\begin{aligned}P[Y=30] &= P[Y=30 \mid D=15]P[D=15]+P[Y=30 \mid D=20]P[D=20]+\ P[Y=30 \mid D=25]P[D=25]\\ &= (0.2)(0.2)+(0.6)(0.5)+(0.7)(0.3)=0.04+0.30+0.21=0.55\end{aligned}$$

The probability that the dense layer is no more than 20' down and that piles must be driven exactly 25' to resistance can be calculated by the corollary of the total probability theorem:

$$\begin{aligned}P[(Y=25)\cap(D=15\cup D=20)] &= P[(Y=25)\cap(D=15)]+P[(Y=25)\cap(D=20)]\\ &= P[Y=25 \mid D=15]P[D=15]+P[Y=25 \mid D=20]P[D=20]\\ &= (0.6)(0.2)+(0.3)(0.5)=0.12+0.15=0.27\end{aligned}$$

The *joint* and *marginal* probabilities for Y and D can be computed by applying the total probability theorem in two steps:

$$P[Y=y_i\cap D=d_j]=P[Y=y_i \mid D=d_j]P[D=d_j]$$
$$P[Y=y_i]=\sum_j P[Y=y_i\cap D=d_j]$$
$$P[D=d_j]=\sum_i P[Y=y_i\cap D=d_j]$$

Joint $P[Y,D]$	$y_1=20$	$y_2=25$	$y_3=30$	$y_4=35$	Marginal $P[D]$
$d_1=15$	0.04	0.12	0.04	0	0.2
$d_2=20$	0	0.15	0.3	0.05	0.5
$d_3=25$	0	0.03	0.21	0.06	0.3
Marginal $P[Y]$	0.04	0.3	0.55	0.11	$\sum=1.00$

© 2024 *Photios G. Ioannou*

Bayes Theorem

Let $B_1, B_2, \ldots, B_n$ be a collection of mutually exclusive and collectively exhaustive events; then

$$P[B_i \mid A] = \frac{P[B_i \cap A]}{P[A]} = \frac{P[A \mid B_i]P[B_i]}{\sum_{i=1}^{n} P[A \mid B_i]P[B_i]}$$

$P[B_i \mid A]$ = *posterior* probability of B_i; i.e., *after* knowing that event *A* has occurred

$P[B_i]$ = *prior* probability of B_i; i.e., *before* knowing that event *A* has occurred

$P[A \mid B_i]$ = *reliability* of the test = *likelihood* of observing *A,* given that the true state is B_i

Using Bayes theorem, we can update the probabilities associated with the possible states of nature B_i based on new evidence *A* from a test of given reliability. Notice that event *A* can be a single test observation, or the entire set of multiple test observations.

Before	**Test**	**After**
Prior probability: $P[B_i]$	New Evidence: Event *A* With Likelihood: $P[A \mid B_i]$	Posterior probability: $P[B_i \mid A]$

Example: Prediction of subsurface ground conditions:

S = {*granite*, limestone, *stiff clay*} = {B_1, B_2, B_3}

Prior probabilities of true ground state: $P[B_1] = 0.5$, $P[B_2] = 0.4$, $P[B_3] = 0.1$

$P[A_j \mid B_i]$	Test Result, A_j			
True ground state, B_i	A_1, favors *granite*	A_2, favors *limestone*	A_3, favors *stiff clay*	$\sum_j P[A_j \mid B_i]$
B_1 = granite	0.7	0.2	0.1	1.0
B_2 = limestone	0.2	0.6	0.2	1.0
B_3 = stiff clay	0.2	0.3	0.5	1.0

Assume that we perform one test and we observe sample outcome A_2, which favors limestone. In this case, the posterior probabilities of the ground states are:

$$P[B_1 \mid A_2] = \frac{P[A_2 \mid B_1]P[B_1]}{P[A_2]} = \frac{P[A_2 \mid B_1]P[B_1]}{P[A_2 \mid B_1]P[B_1] + P[A_2 \mid B_2]P[B_2] + P[A_2 \mid B_3]P[B_3]}$$

$$= \frac{0.2 \times 0.5}{0.2 \times 0.5 + 0.6 \times 0.4 + 0.3 \times 0.1} = \frac{0.1}{0.1 + 0.24 + 0.03} = \frac{0.1}{0.37} = 0.27$$

$$P[B_2 \mid A_2] = \frac{P[A_2 \mid B_2]P[B_2]}{P[A_2]} = \frac{0.6 \times 0.4}{0.37} = 0.65$$

$$P[B_3 \mid A_2] = \frac{P[A_2 \mid B_3]P[B_3]}{P[A_2]} = \frac{0.3 \times 0.1}{0.37} = 0.08$$

Bayes Theorem vs. Subjective Processing of Evidence

Feedback for Inference Study

Judy Shapiro
5263 Institute for Social Research, University of Michigan
747-3650

About 10 years ago, two psychologists named Daniel Kahneman and Amos Tversky gave people the following word problem to solve:

> A cab was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:
>
> (a) 85% of the cabs in the city are Green and 15% are Blue
>
> (b) A witness identified the cab as Blue. The court tested the reliability of the witness under the same conditions that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time.
>
> What is the probability that the cab involved in the accident was Blue?

Very few people were able to answer this problem correctly. The correct answer turns out to be 12/29 = 41%. You can figure out the answer to this problem like this: Let's say the witness was hanging around on the corner, watching cabs go by. For each cab that goes by, what is the probability that the witness will say it is a Blue cab? Well, there is a 15% chance that any given Cab is Blue, and if the Cab really is Blue, there then is an 80% chance that the witness will properly identify the color, which gives (15%)*(80%) = a 12% chance that the cab was Blue and the witness said it was Blue. But the witness might say "That's a Blue cab" even when the cab is Green. The chance that this will happen is (85%)*(20%) = a 17% chance that the cab was really Green, but the witness said it was Blue. The probability that the witness is correct when he says that a cab is Blue is equal to (the probability that a randomly chosen cab is Blue and the witness say it is Blue) divided by (the probability that a randomly chosen cab is Blue and the witness says it is Blue + the probability that a randomly chosen cab is Green and the witness days it is Blue.) This equals 12%/(12%+17%), which is approximately equal to 41%. If this explanation is too difficult for you to understand, don't worry. Like I said, very few people can solve this problem in their heads. And you don't need to understand why the explanation works in order to understand this experiment.

The psychologists who originally gave this problem concluded that people just aren't very good at problems of this sort.

Thus, we should not rely on subjective judgment to update probabilities based on new evidence. It is much more reliable to use Bayes theorem instead.

Events and Random Variables

A variable is a symbol, such as X, which may take any value from some set of numbers.

There are many definitions of a *random variable*. Some of the most straightforward are:

- "*A random variable is a numerical variable whose specific value cannot be predicted with certainty before an experiment*" (Benjamin & Cornell).
- Sometimes, it helps to think of a random variable as a cell in a spreadsheet whose value is not known with certainty before an experiment. To visualize this, we can enter the formula =RAND() into an Excel cell and then press the key F9 (recalculate). Each time F9 is pressed a new "experiment" takes place and the cell (the random variable) takes on a new value (the outcome of the experiment).

In design, an engineer selects or assumes a value for each design variable. The engineer will then specify values for variables in order that the constructed facility will have the properties required in its design. However, the *exact actual value* of most variables is not definite until construction is complete and is usually not known even then.

For example, a structural engineer chooses a concrete strength, a steel strength, an amount of cover over the steel, and a depth of each member. Each of these is specified in the contract documents. The concrete supplier prepares a concrete mix to satisfy the specification. A steel supplier furnishes steel to meet the specification. The rebar placers set the steel with the plans in hand and the formwork carpenters construct the forms to meet the specified dimensions. The actual values of these variables (concrete strength, steel strength, cover and depth) are those which occur in the final construction. And their values vary, by chance, from place to place within the construction. Because their actual values are not definite until construction and they occur by chance, these variables are said to exhibit a random behavior and can be represented by random variables.

In engineering, there are countless random variables. The hardness of subsurface rock, its amount, its depth, and its porosity are all random variables. The quantity of concrete required in a building, its strength at any point, the strength of a test concrete cylinder, and the amount of sand in it are random and so are random variables.

Depending on the values it can take, a random variable may be *discrete* or *continuous*. Two or more random variables may be combined to form another. For example, the average strength of a group of three concrete test cylinders is also a random variable, its value determined by averaging the values of the three cylinders, each a random variable.

In practice, it may be impossible to determine the actual value of every random variable. The strength of concrete at each point in a slab is a random variable; however, we have no way of determining its actual value at each and every point. This does not mean that concrete strength at a particular point is variable (i.e., can change) *after* construction (i.e., when concrete has achieved its long-term strength). It has a particular strength, which even though unknown and indeterminable, is nevertheless definite. If it were 3,486 psi at a point, we would not know it was 3,486 psi nor could we determine that it was 3,486 psi, but nevertheless, it would be 3,486 psi.

© 2024 *Photios G. Ioannou*

We must distinguish in our thinking between

- the true state of nature, and
- our knowledge (or information) about the state of nature.

These are too often confused. A person will assume a value for a random variable, make decisions based upon that assumed value, and then be surprised that the assumed value does not occur and that a decision was in error.

For example, a person standing on the surface of the ground does not usually know the subsurface ground conditions beneath him. He may decide there is a high probability of a firm clay layer below him and base all design and construction decisions upon an assumed capacity of 1.5 TSF at a depth of 20 ft.

Clearly, this assumption does not make it so. The assumption describes one's knowledge of the state of nature; it does not necessarily accurately describe the true state with certainty (as would be the case for a clairvoyant) and certainly does not create that state (as would be the case for a wizard). The soil will not necessarily have an exact capacity of 1.5 TSF at a depth of 20 ft, although hopefully, it will be close. But it may be far different from that which is assumed, and all decisions based upon the assumption could prove wrong and lead to failure.

© 2024 *Photios G. Ioannou*

Random Variables

Random variables are used to describe probabilisitc processes that result in numerical outcomes.
Random variables are usually shown with capital letters from the end of the alphabet: X, Y, Z
Basic event: $a \le X \le b$
$P[X = x]$ = Probability that X takes on the value x, $0 \le P[X = x] \le 1$

Random variables can be discrete or continuous :

Discrete : S is enumerable

Toss of a coin : $S = \{0,1\}$

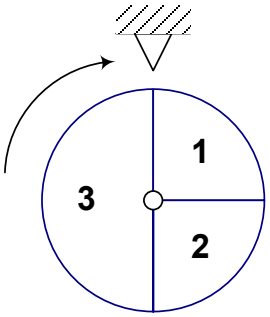
Toss of a die : $S = \{1,2,3,4,5,6\}$

Number of experiments to the first success : $S = \{1,2,3,\cdots\}$

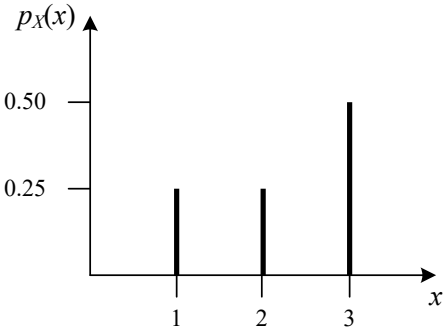
Continuous : S is a subset of the set of real numbers $\Re$

Activity Duration : $S = \{x : a \le x \le b\}$

Time to next arrival : $S = \{x : 0 \le x \le \infty\}$



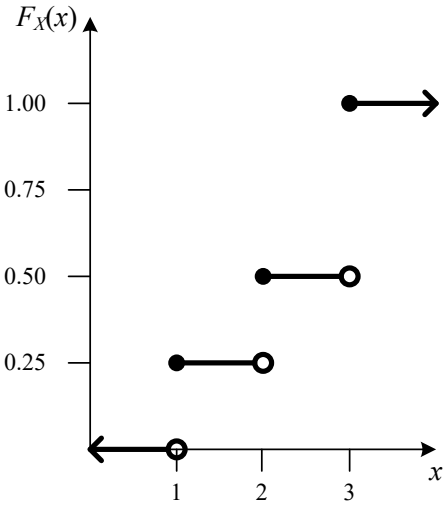
Example Of A Discrete Random Variable



Probability Mass Function (PMF) $= p_X(x) = P[X = x]$

$$P[X = x] = p_X(x) = \begin{cases} 0.25 & x = 1 \\ 0.25 & x = 2 \\ 0.5 & x = 3 \\ 0 & \text{otherwise} \end{cases}$$

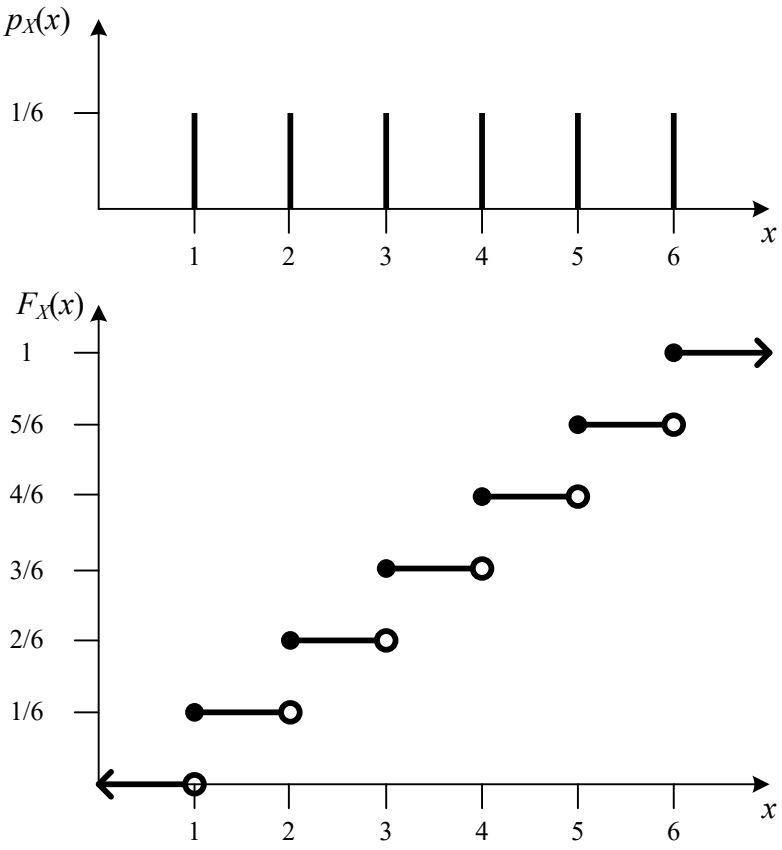
$0 \le p_X(x) \le 1$ and $\sum_{\text{all } x} p_X(x) = 1$



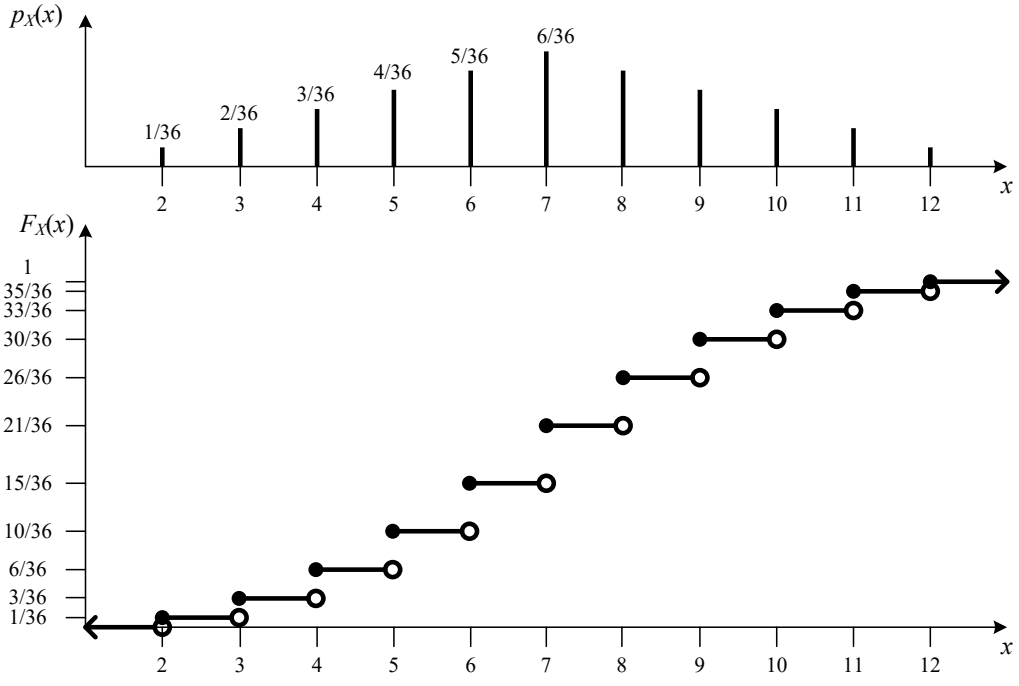
Cumulative Distribution Function (CDF) $= F_X(x) = P[X \le x]$

$$F_X(x) = P[X \le x] = \begin{cases} 0 & x < 1 \\ 0.25 & 1 \le x < 2 \\ 0.5 & 2 \le x < 3 \\ 1 & 3 \le \text{x} \end{cases}$$

$0 \le F_X(x) \le 1$, $F_X(x)$ is a never - decreasing function



Probability Distribution for Single Die



Probability Distribution for Sum of Two Dice

The Bernoulli Random Variable

The Bernoulli random variable, X, is arguably the simplest random variable. It is a binary variable that takes on only the two values x=0, 1. The value 1 is typically called a *success*, and the value 0 a *failure*. They may also represent other binary events, such as *heads* or *tails*.

The probability mass function for the Bernoulli variable is:

$$p_X(x) = \begin{cases} p & \text{for } x = 1 \\ 1-p & \text{for } x = 0 \\ 0 & \text{otherwise} \end{cases}$$

Alternatively, it may be written as:

$$p_X(x) = p^x (1-p)^{1-x} \quad \text{for } x = 0,1$$

The Geometric Distribution

The Geometric random variable, N, is based on the Bernoulli random variable X, and is defined as the number of Bernoulli experiments to the first success. Its sample space is

$$S = \{1, 2, 3, 4, \ldots\}$$

The probability mass function for the number of Bernoulli experiments to the first success, N, can be computed as follows:

$$\begin{aligned} P[N = n] &= P[(n-1 \text{ failures}) \cap (n^{th} \text{experiment a sucess})] \\ &= (1-p)(1-p)\cdots(1-p)p = (1-p)^{n-1} p \end{aligned}$$

Thus, the Geometric probability mass function (PMF) for the random variable N is:

$$p_N(n) = G(p,n) = p(1-p)^{n-1} \qquad \text{for } n = 1, 2, 3, \ldots$$

The cumulative distribution function (CDF) for the random variable N is given by:

$$F_N(n) = P[N \le n] = \sum_{i=1}^{n} p_N(i) = \sum_{i=1}^{n} p(1-p)^{i-1} = p\{(1-p)^0 + (1-p)^1 + \cdots + (1-p)^{n-1}\}$$

Multiplying both sides by $(1-p)$ gives:

$$(1-p)F_N(n) = p\{(1-p)^1 + (1-p)^2 + \cdots + (1-p)^n\}$$

Taking the difference of the last two expressions:

$$F_N(n) - (1-p)F_N(n) = p\{(1-p)^0 - (1-p)^n\}$$

$$pF_N(n) = p\{1-(1-p)^n\}$$

$$F_N(n) = 1-(1-p)^n \qquad n = 1, 2, 3, \ldots$$

The Binomial Distribution

The Binomial random variable, Y, is based on the Bernoulli random variable X. It is defined as:

$$Y = \text{the number of successes in } n \text{ Bernoulli experiments.}$$

Its sample space is

$$S = \{0, 1, 2, \ldots, n\}$$

Its probability mass function can be computed as follows. The probability of any *particular* sequence of outcomes that leads to exactly y successes and (n-y) failures is:

$$p^y (1-p)^{n-y}$$

However, there are many sequences of outcomes (such as, 100111…0) that lead to exactly y successes (*i.e.*, 1's) and (n-y) failures (*i.e.*, 0's). The number of ways we can select y experiments out of n experiments to be the successes is given by the number of *combinations* of n items taken y at a time:

$$\binom{n}{y} = \binom{n}{n-y} = \frac{n!}{y!(n-y)!}$$

Thus, the Binomial probability mass function is:

$$p_Y(y) = \binom{n}{y} p^y (1-p)^{n-y} \quad \text{where } y = 0, 1, \ldots, n$$

In shorthand notation, we write:

$$Y \sim B(n, p)$$
$$Y \sim Bin(n, p)$$

Another way to define the random variable Y is as the sum of n independent and identically distributed (iid) Bernoulli random variables X_i:

$$Y = \sum_{i=1}^{n} X_i$$

This definition is useful for determining the distribution of the sum of two independent random variables each following a Binomial PMF, with the same p. The sum is then a random variable that also follows a Binomial distribution with the same p.

Specifically,

If $Y_1 \sim B(n_1, p)$ and $Y_2 \sim B(n_2, p)$ and Y_1, Y_2 are independent, then

$$Y = Y_1 + Y_2 = \sum_{i=1}^{n_1} X_i + \sum_{j=1}^{n_2} X_j = \sum_{i=1}^{n_1} X_i + \sum_{j=n_1+1}^{n_1+n_2} X_j = \sum_{i=1}^{n_1+n_2} X_i \sim B(n_1 + n_2, p)$$

Jointly Distributed Discrete Random Variables

The *joint probability mass function* of X and Y is given by:

$$p_{X,Y}(x,y) = P[X = x \cap Y = y]$$

The joint probability mass function is a complete description of the random variables X and Y.

We can find the *marginal* PMF of X by using the total probability theorem:

$$p_X(x) = P[X = x] = \sum_{\text{all } y} p_{X,Y}(x,y)$$

Similarly, we find the *marginal* PMF of Y:

$$p_Y(y) = P[Y = y] = \sum_{\text{all } x} p_{X,Y}(x,y)$$

We can find the *conditional* PMF of X given Y, from the definition of conditional probability:

$$p_{X|Y}(x,y) = P[X = x \mid Y = y] = \frac{P[X = x \cap Y = y]}{P[Y = y]} = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

Similarly, we can find the *conditional* PMF of Y given X:

$$p_{Y|X}(x,y) = P[Y = y \mid X = x] = \frac{P[X = x \cap Y = y]}{P[X = x]} = \frac{p_{X,Y}(x,y)}{p_X(x)}$$

The joint cumulative distribution function of X and Y and some of its properties are shown below:

$$F_{X,Y}(x,y) = P[X \le x \cap Y \le y] = \sum_{\text{all } x_i \le x} \sum_{\text{all } y_i \le y} p_{X,Y}(x,y)$$

$$F_{X,Y}(x,\infty) = F_X(x), \quad F_{X,Y}(\infty,y) = F_Y(y), \quad F_{X,Y}(\infty,\infty) = 1, \quad F_{X,Y}(-\infty,-\infty) = 0$$

Independent Random Variables

The definition of *independent random variables* stems directly from the definition of independent events. Thus, the two random variables X and Y are independent if any one of the following equivalent conditions is true:

$p_{X,Y}(x,y) = p_X(x)\,p_Y(y)$ joint PMF = product of marginal PMFs

$p_{X|Y}(x,y) = p_X(x)$ conditional PMF = marginal PMF

$p_{Y|X}(x,y) = p_Y(y)$ conditional PMF = marginal PMF

Example: Jointly Distributed Discrete Random Variables

The jointly distributed random variables *D* and *Y* are defined as follows:

D = the depth of a stratum of dense clay
Y = depth to required resistance for driven piles

The joint probability mass function for *D* and *Y* is shown in the table below:

Joint $p_{D,Y}(d,y)$	y				Marginal $p_D(d)$
d	20	25	30	35	
15	0.04	0.12	0.04	0.00	0.20
20	0.00	0.15	0.30	0.05	0.50
25	0.00	0.03	0.21	0.06	0.30
Marginal $p_Y(y)$	0.04	0.30	0.55	0.11	$\sum = 1.00$

The marginal PMF of *D* is found by adding the joint $p_{D,Y}(d,y)$ across each row.
Similarly, the marginal PMF of *Y* is found by adding the joint $p_{D,Y}(d,y)$ down each column.

The conditional PMF of *Y* given *D* is found by dividing each column of the joint $p_{D,Y}(d,y)$ by the marginal PMF of *D* (the last column):

Conditional $p_{Y\|D}(d,y)$	y			
d	20	25	30	35
15	0.2	0.6	0.2	0
20	0	0.3	0.6	0.1
25	0	0.1	0.7	0.2

The conditional PMF of *D* given *Y* is found by dividing each row of the joint $p_{D,Y}(d,y)$ by the marginal PMF of *Y* (the last row):

Conditional $p_{D\|Y}(d,y)$	y			
d	20	25	30	35
15	1	0.4	0.072727	0
20	0	0.5	0.545455	0.454545
25	0	0.1	0.381818	0.545455

© 2024 *Photios G. Ioannou*

Example: Bernoulli Experiments and Jointly Distributed Random Variables

This is a continuation of the previous example where a superintendent requests of the union business agent that 5 carpenters be assigned for a Friday morning.

The random variable N (the number of carpenters actually assigned) can range from 0 to 5. The probability mass function for N is given by the following table:

n =	0	1	2	3	4	5
$P[N=n]$ =	0.15	0.05	0.05	0.15	0.2	0.4

In this example, we consider that the carpenters' union includes two types of workers: *journeymen* and *apprentices*. Hence, the *joint sample space* for the *number* of journeymen and the *number* of apprentices actually assigned is triangular and would look as follows:

$$\begin{array}{cccccc} E_{0,0} & E_{0,1} & E_{0,2} & E_{0,3} & E_{0,4} & E_{0,5} \\ E_{1,0} & E_{1,1} & E_{1,2} & E_{1,3} & E_{1,4} & \\ E_{2,0} & E_{2,1} & E_{2,2} & E_{2,3} & & \\ E_{3,0} & E_{3,1} & E_{3,2} & & & \\ E_{4,0} & E_{4,1} & & & & \\ E_{5,0} & & & & & \end{array}$$

Each simple event $E_{j,a}$ represents the case where j = journeymen and a = apprentices.

Each simple event $N = n$ can now be represented by a set of simple events $E_{j,a}$ such that $j + a = n$. For example, the event $N = 3$ includes all events $E_{j,a}$ for which $j + a = 3$, all four of which lie on the same diagonal from lower left to upper right: $E_{3,0}, E_{2,1}, E_{1,2}, E_{0,3}$.

If 1/5 of the members of the carpenters union are apprentices, and the status of each person assigned is independent of the status of any other, then the situation is similar to performing an random number of Bernoulli experiments N, where each experiment can be a success (journeyman) with probability p = 4/5 or failure (apprentice) with probability 1- p =1/5.

The sample points in the triangular matrix above are mutually exclusive events. For example, event $E_{2,1}$ is the assignment of exactly 3 carpenters (experiments), of which there are exactly 2 journeymen (successes) and 1 apprentice (failure). Moreover, each event $E_{j,a}$ is made up of several more detailed sample points. For example, event $E_{2,1}$ is made up of the three events *JJA*, *JAJ*, *AJJ*, where *JJA* is the event where "of the three workers assigned the first is a journeyman, the second is a journeyman, and the third is an apprentice."

Thus, the triangular matrix of events $E_{j,a}$ can be rewritten as follows:

$^{(1)}E_{0,0}$	$^{(1)}E_{0,1}$	$^{(1)}E_{0,2}$	$^{(1)}E_{0,3}$	$^{(1)}E_{0,4}$	$^{(1)}E_{0,5}$
	A	AA	AAA	$AAAA$	$AAAAA$
$^{(1)}E_{1,0}$	$^{(2)}E_{1,1}$	$^{(3)}E_{1,2}$	$^{(4)}E_{1,3}$	$^{(5)}E_{1,4}$	
J	JA, AJ	JAA, AJA, AAJ	$4(JAAA)$	$5(JAAAA)$	
$^{(1)}E_{2,0}$	$^{(3)}E_{2,1}$	$^{(6)}E_{2,2}$	$^{(10)}E_{2,3}$		
JJ	JJA, JAJ, AJJ	$JJAA, JAJA, JAAJ,$ $AJJA, AJAJ, AAJJ$ $(P[E_{2,2}] = 6P[J]^2 P[A]^2)$	$10(JJAAA)$, where $10 = \binom{5}{2} = \binom{5}{3}$		
$^{(1)}E_{3,0}$	$^{(4)}E_{3,1}$	$^{(10)}E_{3,2}$			
JJJ	$4(JJJA)$	$10(JJJAA)$			
$^{(1)}E_{4,0}$	$^{(5)}E_{4,1}$				
$JJJJ$	$5(JJJJA)$				
$^{(1)}E_{5,0}$					
$JJJJJ$					

$^{(k)}E_{j,a}$: 21 events

j = no. of journeymen

a = no. of apprentices

Each event $^{(k)}E_{j,a}$ contains *k sequences* (such as JAA) where the type (J vs. A) and the order of worker arrival matters. There are 63 such *sequences* (compound events).

The easiest way to calculate the probability of an event $E_{j,a}$, is to recognize that the random variable J given that $J + A = N = n$ follows a Binomial distribution.

For example, for event $E_{2,1}$ we have $J = 2$ and $A = 1$, and thus $J + A = N = 3$. If we let J be a *success*, then we seek the probability of having exactly 2 successes out of 3 experiments multiplied by the probability of having exactly 3 experiments:

$$\begin{aligned} P[E_{2,1}] &= P[(J=2) \cap (A=1)] \\ &= P[(J=2) \cap (N=3)] \\ &= P[J=2 \mid N=3]\ \ P[N=3] \\ &= \binom{3}{2} p^2\ \ (1-p)^{3-2} P[N=3] \\ &= 3 \times (\frac{4}{5})^2 \times (\frac{1}{5})^1 \times 0.15 = 3 \times \frac{16}{125} \times 0.15 = 0.0576 \end{aligned}$$

The following table shows the probabilities for all sample points $P[E_{j,a}]$.

$P[E_{j,a}]$	A = no. of apprentices						
J = no. of journeymen	0	1	2	3	4	5	$\sum \rightarrow P[J=j]$
0	0.150	0.010	0.002	0.0012	0.0003	0.0001	0.164
1	0.040	0.016	0.014	0.005	0.003		0.078
2	0.032	0.058	0.031	0.020			0.140
3	0.077	0.082	0.082				0.241
4	0.082	0.164					0.246
5	0.131						0.131
$\sum \rightarrow P[A=a]$	0.512	0.330	0.129	0.026	0.006	0.001	$\sum = 1.00$

If we wish to know the probability of having exactly 4 journeymen carpenters assigned:

$$
\begin{aligned}
P[J=4] &= P[(J=4 \cap A=0) \cup (J=4 \cap A=1)] \\
&= P[E_{4,0} \cup E_{4,1}] = P[E_{4,0}] + P[E_{4,1}] \\
P[E_{4,0}] &= P[J=4 \mid N=4]P[N=4] \\
&= (\frac{4}{5})^4 \times 0.2 = \frac{256}{625} \times 0.2 = 0.082 \\
P[E_{4,1}] &= P[J=4 \mid N=5]P[N=5] \\
&= \binom{5}{4}\left(\frac{4}{5}\right)^4\left(\frac{1}{5}\right) \times 0.4 = 0.410 \times 0.4 = 0.164 \\
P[J=4] &= 0.082 + 0.164 = 0.246
\end{aligned}
$$

We see from this tabulation of sample point probabilities that a summation of rows will give the probabilities for the number of apprentices, *A*, whereas a summation of columns gives probabilities for journeymen, *J*. The summation of all sample points equals the certain event with a probability of one.

Notice that if we add up the probabilities in each diagonal of the above table, from lower left to upper right, we obtain the probability that a particular number of workers are assigned. The results are shown in the table below and match the original data for this example:

$n = j + a$	0	1	2	3	4	5	
$P[J+A=n] = P[N=n] =$	0.15	0.05	0.05	0.15	0.2	0.4	$\sum = 1.00$

Note: In practice, the number of assigned apprentices may <u>not</u> be independent of the number of journeymen assigned. In such case, the above table cannot be computed without other conditional probabilities not given here.

Example: Non-Bernoulli Experiments on Jointly Distributed Random Variables

Let us now assume that the labor contract calls for one apprentice, if available, whenever four or more carpenters are assigned to a job, and that we are given the following:

$$P[A=1 \mid N=4]=0.2$$
$$P[A=1 \mid N=5]=0.6$$

The sample points for the random variable J (the number of journeymen) are then these:

$J=0$	$E_{0,0}$	$P[J=0]=P[N=0]=0.15$ (given)
$J=1$	$E_{1,0}$	$P[J=1]=P[N=1]=0.05$ (given)
$J=2$	$E_{2,0}$	$P[J=2]=P[N=2]=0.05$ (given)
$J=3$	$E_{3,0}$, $E_{3,1}$	$P[J=3]=0.19$ (see below)
$J=4$	$E_{4,0}$, $E_{4,1}$	$P[J=4]=0.40$ (see below)
$J=5$	$E_{5,0}$	$P[J=5]=0.16$ (see below)
		$\sum P[J]=1.00$

The first three probabilities $P[J=j]$ were given as $P[N=n]$.

The probability $P[E_{2,1}]=0$ since no apprentice will be assigned with only 2 journeymen.

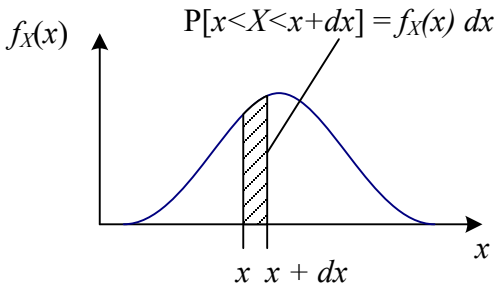
The probabilities for the last three rows are computed using the Total Probability Theorem:

$P[J=3]$	$P[E_{3,0}]=P[N=3]=0.15$ (given) $P[E_{3,1}]=P[J=3\cap A=1]=P[A=1\cap N=4]$ $=P[A=1 \mid N=4]P[N=4] = 0.2 \times 0.2 = 0.04$ $P[J=3] = P[E_{3,0}] + P[E_{3,1}] = 0.15 + 0.04 = 0.19$
$P[J=4]$	$P[E_{4,0}] = P[J=4\cap A=0]=P[A=0\cap N=4]$ $= P[A=0 \mid N=4]P[N=4]$ $= (1-P[A=1 \mid N=4])P[N=4] = (1-0.2)(0.2) = 0.16$ $P[E_{4,1}] = P[J=4\cap A=1]= P[A=1\cap N=5]$ $= P[A=1 \mid N=5]P[N=5]= (0.6)(0.4) = 0.24$ $P[J=4] = P[E_{4,0}] + P[E_{4,1}]=0.16 + 0.24 = 0.40$
$P[J=5]$	$P[E_{5,0}]=P[J=5\cap A=0]=P[A=0\cap N=5]$ $=P[A=0 \mid N=5]P[N=5]$ $= (1-P[A=1 \mid N=5])P[N=5] = (1-0.6)(0.4) = 0.16$ $P[J=5] = P[E_{5,0}]=0.16$

Continuous Random Variables

For a continuous random variable $P[X = x] = 0 \quad \Rightarrow \quad$ no probability mass exists at any point x

Instead, points x are associated with probability density: Probability Density Function (PDF) $= f_X(x)$



$$0 \le f_X(x), \qquad f_X(x) \text{ may be} > 1$$

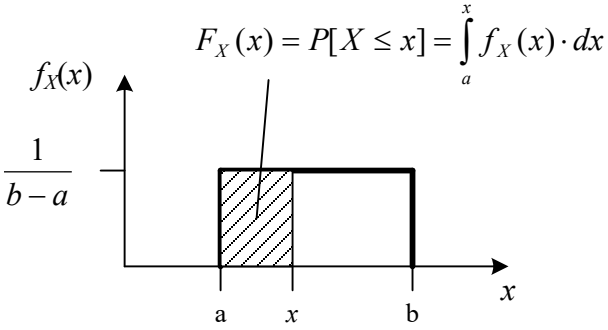
$$P[S] = \int_{-\infty}^{+\infty} f_X(x) \cdot dx = 1$$

$$F_X(x) = P[X \le x] = \int_{-\infty}^{x} f_X(x) \cdot dx$$

$$P[a \le X \le b] = \int_{a}^{b} f_X(x) \cdot dx = F_X(b) - F_X(a)$$

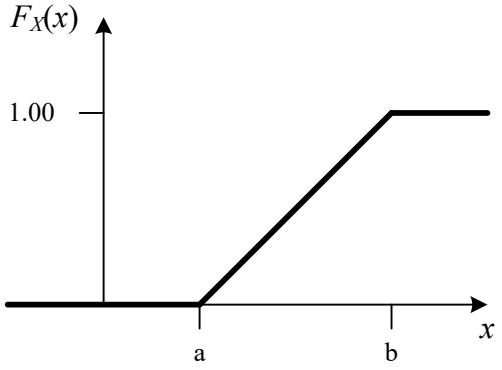
$$f_X(x) = \frac{dF_X(x)}{dx}$$

The Uniform Distribution



Uniform PDF

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$



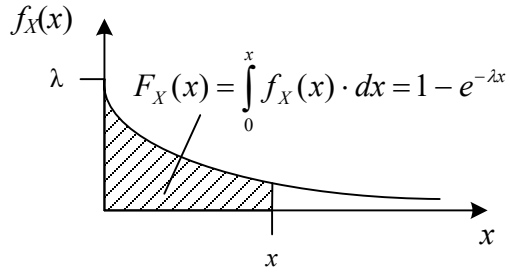
CDF of Uniform PDF

For $a \le x \le b$

$$F_X(x) = \int_a^x \frac{1}{b-a} \cdot dx = \left[\frac{x}{b-a} \right]_a^x = \frac{x-a}{b-a}$$

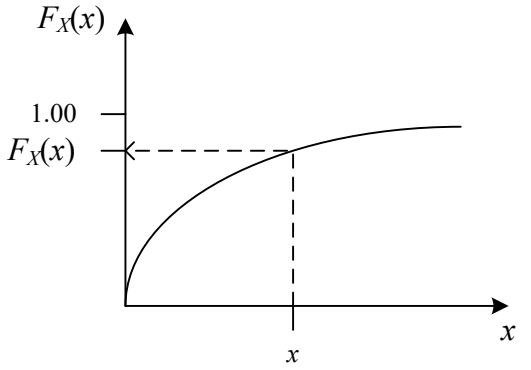
$$F_X(x) = \begin{cases} 0 & x \le a \\ \dfrac{x-a}{b-a} & a \le x \le b \\ 1 & b \le x \end{cases}$$

The Exponential Distribution



Exponential PDF

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & 0 \le x \\ 0 & \text{otherwise} \end{cases}$$



Exponential CDF

$$F_X(x) = \int_0^x \lambda e^{-\lambda x} dx = [-e^{-\lambda x}]_0^x$$

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & 0 \le x \\ 0 & \text{otherwise} \end{cases}$$

Jointly Distributed Continuous Random Variables

The *joint probability density function* of X and Y is given by $f_{X,Y}(x,y)$, where:

$$f_{X,Y}(x,y)\,dx\,dy = P[x \le X \le x+dx \cap y \le Y \le y+dy]$$

The joint probability density function is a complete description of the random variables X and Y.

The probability of the basic event A is given by,

$$P[A] = P[a_1 \le X \le b_1 \cap a_2 \le Y \le b_2] = \int_{x=a_1}^{b_1} \int_{y=a_2}^{b_2} f_{X,Y}(x,y)\,dx\,dy$$

From the Total Probability Theorem, we can find the *marginal* PDFs for X and for Y:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\;dy \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\;dx$$

The *conditional* PDF of X given Y:

$$f_{X|Y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

and the *conditional* PDF of Y given X:

$$f_{Y|X}(x,y) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

The cumulative distribution function defined below has the following properties:

$$F_{X,Y}(x,y) = P[X \le x \cap Y \le y] = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(x,y)\,dx\,dy$$

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$$

$$F_{X,Y}(x,\infty) = F_X(x), \quad F_{X,Y}(\infty,y) = F_Y(y), \quad F_{X,Y}(\infty,\infty) = 1, \quad F_{X,Y}(-\infty,-\infty) = 0$$

Independent Random Variables

The continuous random variables X and Y are independent if any one of the following equivalent conditions is true:

$$f_{X,Y}(x,y) = f_X(x) f_Y(y)$$

$$f_{X|Y}(x,y) = f_X(x)$$

$$f_{Y|X}(x,y) = f_Y(y)$$

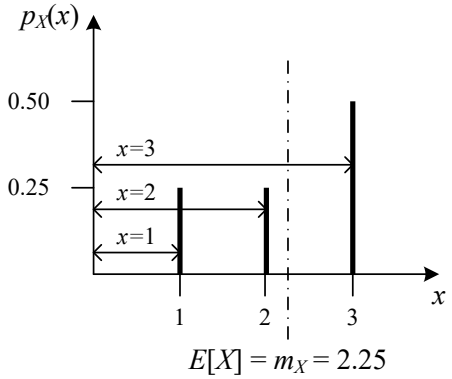
Moments of Discrete Random Variables

Expected Value (First Moment)

Expected Value = Mean = Centroid of probability mass = Limit of long - term average of outcomes

$$E[X] = m_X = \sum_{\text{all } x} x \cdot p_X(x)$$

$$E[X] = m_X = 1 \times 0.25 + 2 \times 0.25 + 3 \times 0.5 = 2.25$$



Variance (Second Centroidal Moment)

Variance = Centroidal moment of inertia of probability mass = Measure of dispersion (variability)

$$Var[X] = \sigma_X^2 = \sum_{\text{all } x} (x - E[X])^2 \cdot p_X(x)$$

$$Var[X] = \sigma_X^2 = 1.25^2 \times 0.25 + 0.25^2 \times 0.25 + 0.75^2 \times 0.5 = 0.6875$$

Theorem of parallel axes

$$Var[X] = \sigma_X^2 = E[(X - E[X])^2] = E[X^2 - 2XE[X] + (E[X])^2] = E[X^2] - 2E[X]E[X] + E[(E[X])^2]$$

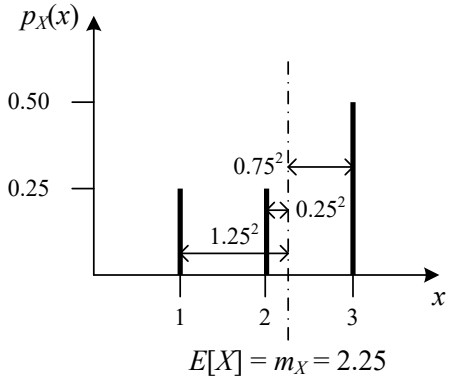
$$= E[X^2] - 2(E[X])^2 + (E[X])^2 = E[X^2] - (E[X])^2$$

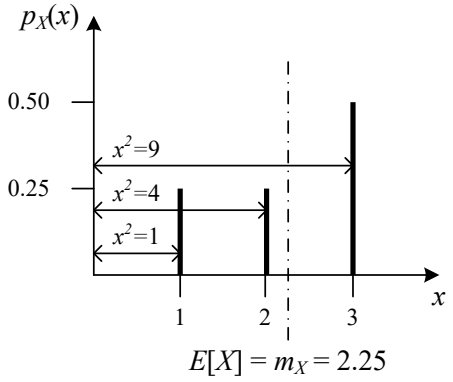
$$E[X^2] = 1^2 \times 0.25 + 2^2 \times 0.25 + 3^2 \times 0.5 = 0.25 + 1.00 + 4.5 = 5.75$$

$$Var[X] = \sigma_X^2 = E[X^2] - (E[X])^2 = 5.75 - 2.25^2 = 0.6875$$

Standard Deviation (has the same units as X)

$$SD[X] = \sigma_X = \sqrt{Var[X]} = \sqrt{0.6875} = 0.83$$





Example: Expected Value of a Function

Often we are interested in the expected value of *Y*, where $Y = g(X)$ is a function of a random variable *X*. It is important to appreciate that in general we cannot just substitute the expected value of *X* into $g(X)$:

$$E[Y] = E[g(X)] \neq g(E[X])$$

As shown in the example below, there are two correct ways to calculate the expected value $E[Y]$. Let

$$p_X(x) = \begin{cases} \frac{1}{2} & \text{for } x = 1,3 \\ 0 & \text{otherwise} \end{cases} \quad \Rightarrow \quad E[X] = m_X = \sum_{\text{all } x} x \cdot p_X(x) = \frac{1}{2} \times 1 + \frac{1}{2} \times 3 = 2$$

The first requires that we determine the probability mass function of *Y*:

Let $Y = g(X) = \frac{1}{X}$

$$p_Y(y) = \begin{cases} \frac{1}{2} & \text{for } y = \frac{1}{3},1 \\ 0 & \text{otherwise} \end{cases} \quad \Rightarrow \quad E[Y] = \frac{1}{2} \times 1 + \frac{1}{2} \times \frac{1}{3} = \frac{2}{3}$$

The second is a direct computation of the expected value of *Y*, as a function $g(X)$. It gives the same result but does not require the probability mass function of *Y*,

$$\begin{aligned} E[Y] &= E[g(X)] \\ &= \sum_{\text{all } x} g(x) \cdot p_X(x) \\ &= g(1)p_X(1) + g(3)p_X(3) \\ &= 1 \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{2} = \frac{2}{3} \end{aligned}$$

Notice, however, that

$$E[Y] = E[g(X)] = \frac{2}{3} \neq g(E[X]) = \frac{1}{E[X]} = \frac{1}{2}$$

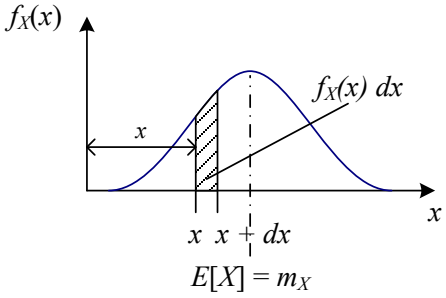
The only exception is when $g(X)$ is a linear function of *X*.

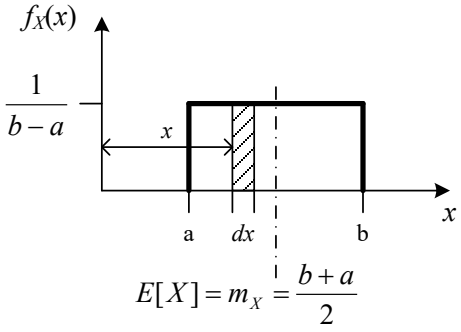
Moments of Continuous Random Variables

Expected Value (First Moment)

Expected Value = Mean = Centroid of area under PDF = Limit of long - term average of outcomes

$$E[X] = m_X = \int_{-\infty}^{+\infty} x \cdot f_X(x)dx \qquad E[X] = m_X = \int_a^b x \cdot \frac{1}{b-a}dx = \left[\frac{x^2}{2(b-a)}\right]_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{b+a}{2}$$

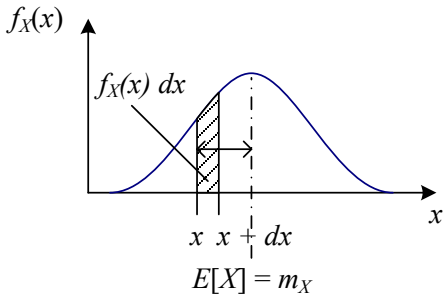


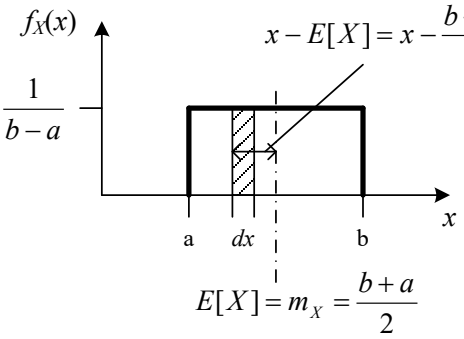


Variance (Second Centroidal Moment)

Variance = Centroidal moment of inertia of area under PDF = Measure of dispersion (variability)

$$Var[X] = \sigma_X^2 = \int_{-\infty}^{+\infty} (x - E[x])^2 \cdot f_X(x)dx \qquad Var[X] = \sigma_X^2 = \int_a^b (x - \frac{b+a}{2})^2 \cdot \frac{1}{b-a}dx$$





$$Var[X] = \sigma_X^2 = E[(x - E[x])^2] = E[X^2] - (E[X])^2$$

$$E[X^2] = \int_a^b \frac{x^2 dx}{b-a} = \left[\frac{x^3}{3(b-a)}\right]_a^b = \frac{b^3-a^3}{3(b-a)} = \frac{a^2+ab+b^2}{3}$$

$$Var[X] = \frac{a^2+ab+b^2}{3} - \frac{(b+a)^2}{4} = \frac{(b-a)^2}{12}$$

Standard Deviation (has the same units as X)

$$SD[X] = \sigma_X = \sqrt{Var[X]} = \frac{(b-a)}{2\sqrt{3}}$$

BASIC PROBABILITY DISTRIBUTIONS

	Discrete Experiments Bernoulli Trials	**Continuous Time** Poisson Arrivals
Random Variable	$p = P$[success in any trial] = mean number of successes per trial	λ = mean rate of arrivals = mean no. of arrivals per unit time
Binomial: Y = *number of successes* in n experiments. *Poisson:* Y = *number of arrivals* in t time.	**Binomial Distribution** $Y \sim B(n, p)$ $p_Y(y) = \binom{n}{y} p^y (1-p)^{n-y}$ where $y = 0,1,...,n$ $E[Y] = np$ $Var[Y] = np(1-p)$	**Poisson Distribution** $Y \sim P(\lambda t)$ $p_Y(y) = \frac{(\lambda t)^y e^{-\lambda t}}{y!}$ where $y = 0,1,2,...$ $E[Y] = Var[Y] = \lambda t = \nu$
Geometric: N = number of experiments to the first success. *Exponential*: T = interarrival time = length of time to the next arrival.	**Geometric Distribution** $N \sim G(p)$ $p_N(n) = (1-p)^{n-1} p \qquad n = 1,2,3,...$ $F_N(n) = 1-(1-p)^n \qquad n = 1,2,3,\ldots$ $E[N] = \frac{1}{p} \qquad Var[N] = \frac{1-p}{p^2}$ Memoryless Property $P[N = n+k \mid N > k] = P[N = n]$	**Exponential Distribution** $T \sim Exp(\lambda)$ $f_T(t) = \lambda e^{-\lambda t} \qquad t \geq 0$ $F_T(t) = 1-e^{-\lambda t} \qquad t \geq 0$ $E[T] = \frac{1}{\lambda} \qquad Var[T] = \frac{1}{\lambda^2}$ Memoryless Property: $P[T > t+k \mid T > k] = P[T > t]$
Neg. Binomial: N_k =number of experiments to the k^{th} success. *Erlang:* T_k = length of time to the k^{th} arrival	**Negative Binomial Distribution** $N_k \sim NB(p, k)$ $p_{N_k}(n) = \binom{n-1}{k-1} p^k (1-p)^{n-k}$ where $n = k, k+1, \ldots$ $E[N_k] = \frac{k}{p} \qquad Var[N_k] = \frac{k(1-p)}{p^2}$ Geometric $G(p) = NB(p,1)$	**Erlang Distribution** $T_k \sim Erl(\lambda, k)$ $f_{T_k}(t) = \frac{\lambda(\lambda t)^{k-1} e^{-\lambda t}}{(k-1)!}$ where $t \geq 0$ $E[T_k] = \frac{k}{\lambda} \qquad Var[T_k] = \frac{k}{\lambda^2}$ Exponential $Exp(\lambda) = Erl(\lambda,1)$

PROPERTIES OF SOME BASIC PROBABILITY DISTRIBUTIONS

The sum of *n* iid (i.e., *Independent and Identically Distributed*) Bernoulli random variables (i.e., with the same *p*), is a random variable that follows a Binomial distribution with the same *p*:

$$\text{If } X_i \text{ are iid and } X_i \sim Bernoulli(p) \qquad i = 1,...,n, \qquad \Rightarrow \qquad Y = \sum_{i=1}^{n} X_i \sim B(n,p)$$

The sum of independent random variables each following a Binomial PMF, with the same *p*, is a random variable that also follows a Binomial distribution with the same *p*:

If $Y_1 \sim B(n_1,p)$ and $Y_2 \sim B(n_2,p)$ and Y_1, Y_2 are independent, then $Y = Y_1 + Y_2 \sim B(n_1+n_2,p)$

The sum of *k* iid random variables each following a Geometric PMF, with the same *p*, is a random variable that follows a Negative Binomial distribution with parameters *k* and *p*:

$$\text{If all } N_j \sim G(p), \text{ and } all\ N_j \text{ are independent} \Rightarrow N_k = \sum_{j=1}^{k} N_j \sim NB(p,k)$$

The sum of independent random variables each following a Negative Binomial PMF, with the same *p*, is a random variable that also follows a Negative Binomial distribution with the same *p*:

If $N_{k_1} \sim NB(p,k_1)$ and $N_{k_2} \sim NB(p,k_2)$ and N_{k_1}, N_{k_2} are independent, then

$$N_{k_1+k_2} = N_{k_1} + N_{k_2} \sim NB(p,k_1+k_2)$$

Bernoulli sampling (with probability of success *p*) on a Poisson process with parameter, λ, yields a Poisson process with parameter λp.

The sum of independent random variables each following a Poisson PMF is a random variable that also follows a Poisson distribution:

$$\text{If } \left\{ \begin{array}{c} Y_1 \sim P(\nu_1) \ and\ Y_2 \sim P(\nu_2) \\ or \\ Y \sim P(\lambda_1 t_1) \ and\ Y_2 \sim P(\lambda_2 t_2) \end{array} \right\} \text{ and } Y = Y_1 + Y_2 \Rightarrow \left\{ \begin{array}{c} Y \sim P(\nu_1+\nu_2) \\ or \\ Y \sim P(\lambda_1 t_1 + \lambda_2 t_2) \end{array} \right\}$$

The sum of *k* iid random variables each following a Exponential PDF, with the same λ, is a random variable that follows an Erlang distribution with parameters *k* and λ:

$$\text{If all } T_j \sim Exp(\lambda), \text{ and } all\ T_j \text{ are independent} \Rightarrow T_k = \sum_{j=1}^{k} T_j \sim Erl(\lambda,k)$$

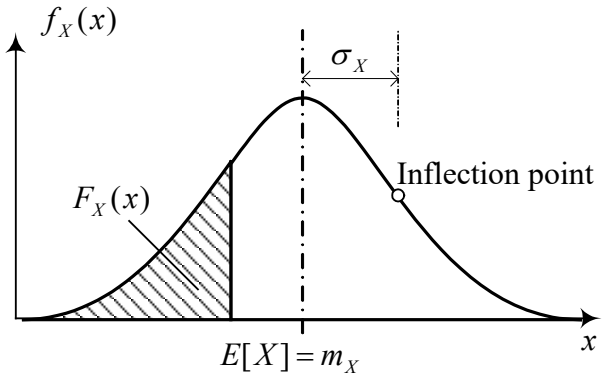
The sum of independent random variables each following an Erlang PDF, with the same λ, is a random variable that also follows an Erlang distribution with the same λ:

If $T_{k_1} \sim Erl(\lambda,k_1)$ and $T_{k_2} \sim Erl(\lambda,k_2)$ and T_{k_1}, T_{k_2} are independent, then

$$T_{k_1+k_2} = T_{k_1} + T_{k_2} \sim Erl(\lambda,k_1+k_2)$$

© 2024 *Photios G. Ioannou*

THE NORMAL DISTRIBUTION

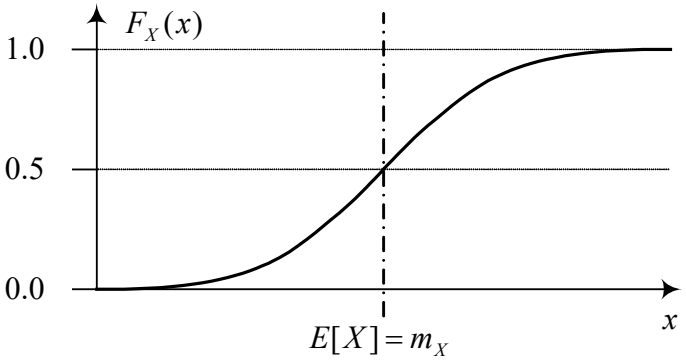


$$X \sim N(m_X, \sigma_X)$$

$$f_X(x) = \frac{1}{\sigma_X \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-m_X}{\sigma_X})^2}$$

$F_X(x)$ does not exist. I.e., there is no function, such that $\frac{dF_X(x)}{dx} = f_X(x)$.

Thus, $F_X(x)$ can only be evaluated using numerical methods for specific values of x, m_X, and σ_X.

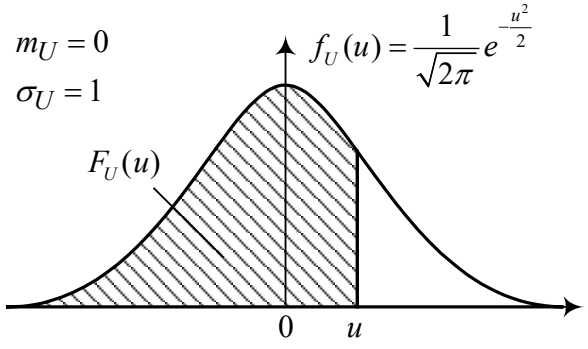


THE STANDARD NORMAL VARIATE

$$U = \frac{X - m_X}{\sigma_X}$$

The CDF $F_U(u)$ is widely available in tables. These tables usually show $F_U(u)$ only for positive values of u, because:

$$F_U(-u) = 1 - F_U(u)$$

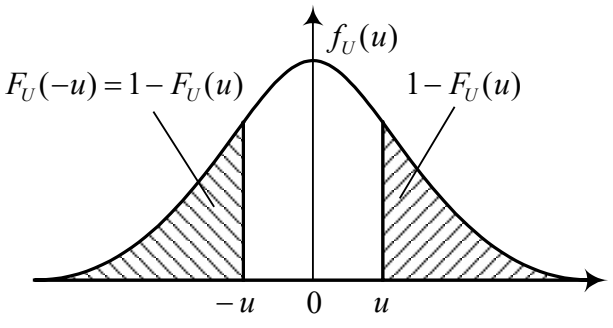


HOW TO CALCULATE $F_X(x)$ FOR ANY m_X and σ_X

$$F_X(x) = P[X \le x] = P[\frac{X - m_X}{\sigma_X} \le \frac{x - m_X}{\sigma_X}] =$$

$$P[U \le \frac{x - m_X}{\sigma_X}] = F_U(\frac{x - m_X}{\sigma_X}) \quad \Rightarrow$$

$$\boxed{F_X(x) = F_U(\frac{x - m_X}{\sigma_X})}$$



Example: $X \sim N(6,4)$ $\quad m_X = 6,\ \sigma_X^2 = 16$

$$F_X(x) = F_U(\frac{x-6}{4})$$

$$F_X(10) = F_U(\frac{10-6}{4}) = F_U(1) = 0.8413$$

$$F_X(-1) = F_U(\frac{-1-6}{4}) = F_U(-1.75) =$$

$$= 1 - F_U(1.75) = 1 - 0.96 = 0.04$$

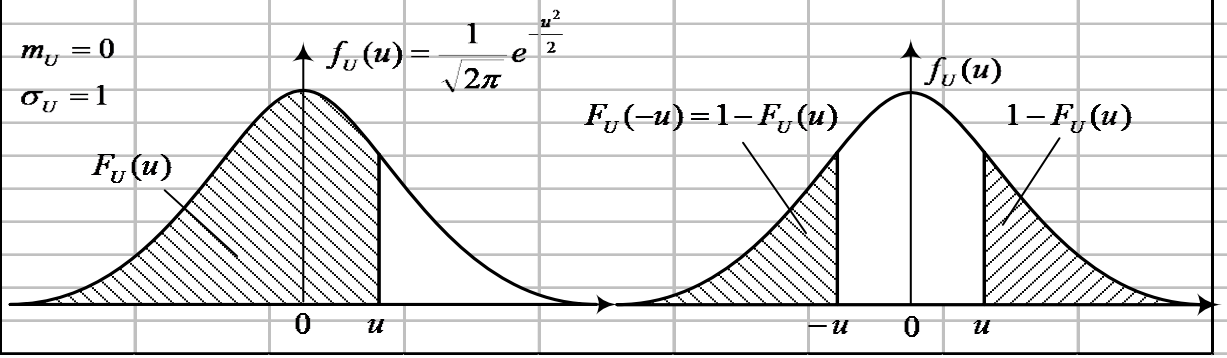
© 2024 *Photios G. Ioannou*

Values of the Standardized Normal Distribution, $U \sim N(0,1)$

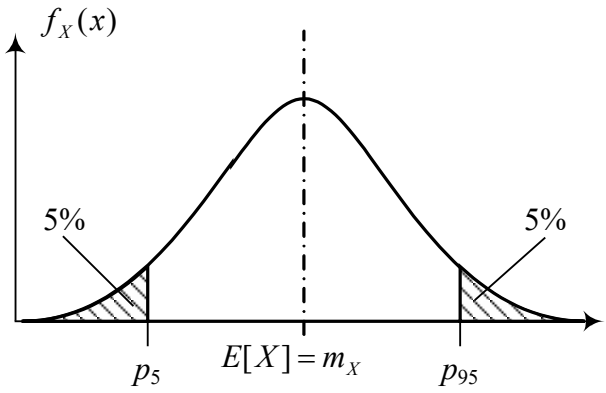
Table gives values of the Cumulative Distribution Function $F_U(u)$, for $u > 0$

For $-u < 0$, apply $F_U(-u) = 1 - F_U(u)$

u	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	Percentiles $F(u)$	Percentiles u
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359	50.0%	0.00000
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753	55.0%	0.12566
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141	60.0%	0.25335
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517	65.0%	0.38532
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879	70.0%	0.52440
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224	75.0%	0.67449
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549	80.0%	0.84162
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852	82.5%	0.93459
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133	85.0%	1.03643
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389	87.5%	1.15035
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621	90.0%	1.28155
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830	90.5%	1.31058
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.90147	91.0%	1.34076
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774	91.5%	1.37220
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189	92.0%	1.40507
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408	92.5%	1.43953
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449	93.0%	1.47579
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327	93.5%	1.51410
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062	94.0%	1.55477
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670	94.5%	1.59819
2.0	0.97725										95.0%	1.64485
2.1	0.98214										95.5%	1.69540
2.2	0.98610										96.0%	1.75069
2.3	0.98928										96.5%	1.81191
2.4	0.99180										97.0%	1.88079
2.5	0.99379										97.5%	1.95996
3.0	0.99865										98.0%	2.05375
3.5	0.999767										98.5%	2.17009
4.0	0.9999683										99.0%	2.32635
4.5	0.9999966										99.25%	2.43238
5.0	0.99999971										99.50%	2.57583
5.5	0.999999981										99.75%	2.80703



PERCENTILES, MEDIAN, MODE



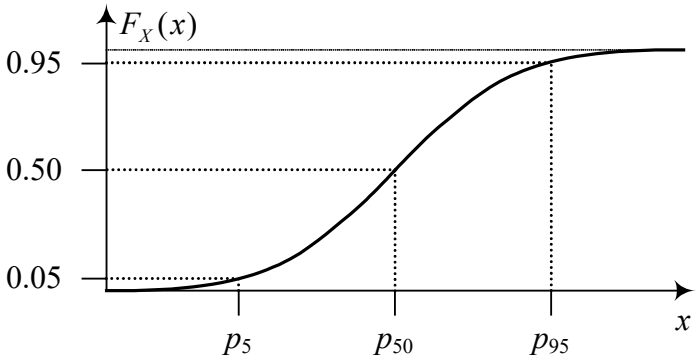
PERCENTILES

The kth percentile, p_k, of a random variable X is that value x such that

$$F_X(p_k) = k\%$$

In PERT we will find the following three percentiles useful for the definition of the Perry and Greig formulas:

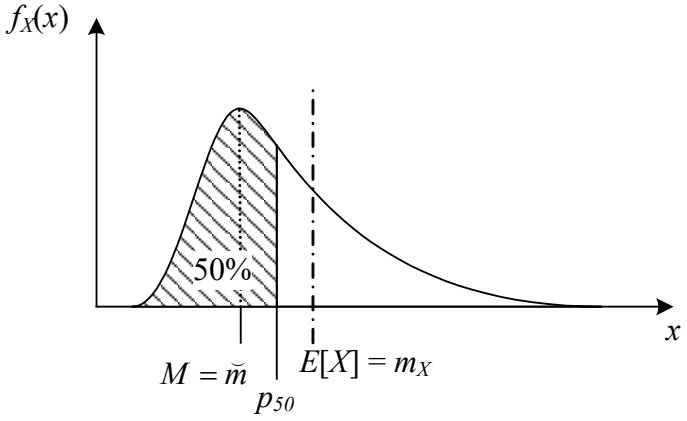
$$p_5 = 5^{th} \text{ percentile}: \quad F_X(p_5) = 5\%$$
$$p_{50} = 50^{th} \text{ percentile}: \quad F_X(p_{50}) = 50\%$$
$$p_{95} = 95^{th} \text{ percentile}: \quad F_X(p_{95}) = 95\%$$



MEDIAN

The median p_{50} of a random variable X divides the area under the PDF into two parts of equal area. By definition, the median is equal to the 50^{th} percentile:

$$F_X(p_{50}) = 50\%$$



MODE

The mode $M = \breve{m}$ of a random variable X is that value of x where the PDF $f_X(x)$ has a maximum:

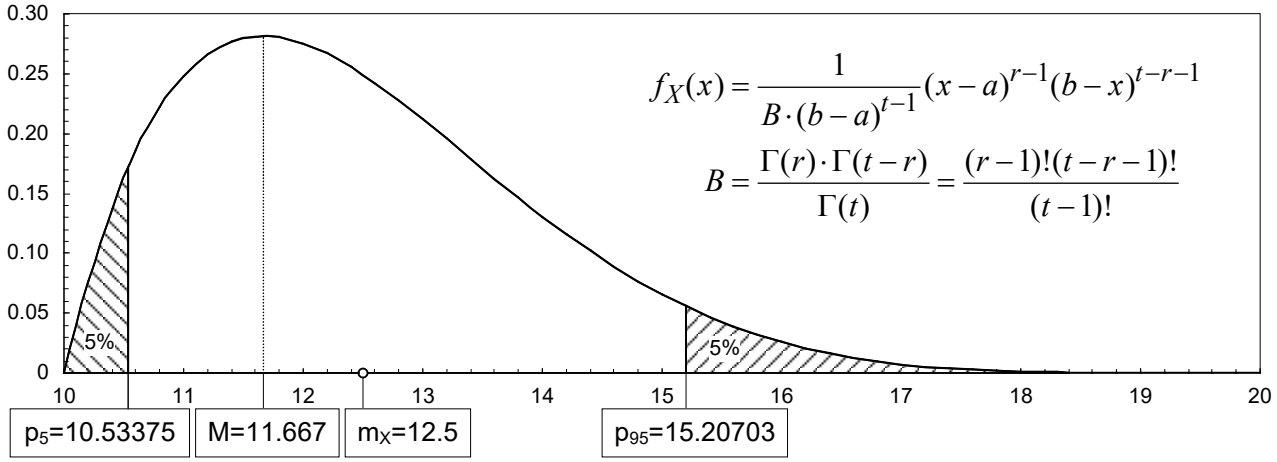
$$\left.\frac{df_X(x)}{dx}\right|_M = 0$$

The mode is also called

"the most likely value of X".

Most distributions have a single mode (i.e., they are unimodal).

APPROXIMATIONS FOR THE MOMENTS OF THE BETA DISTRIBUTION



$$f_X(x) = \frac{1}{B \cdot (b-a)^{t-1}} (x-a)^{r-1} (b-x)^{t-r-1}$$

$$B = \frac{\Gamma(r) \cdot \Gamma(t-r)}{\Gamma(t)} = \frac{(r-1)!(t-r-1)!}{(t-1)!}$$

$$\left.\begin{array}{l} a = 10 \\ b = 20 \\ r = 2 \\ t = 8 \end{array}\right\} \Rightarrow \left\{\begin{array}{l} m_X = a + \dfrac{r}{t}(b-a) = 10 + \dfrac{2}{8}(20-10) = 12.5 \\ \sigma_X^2 = (b-a)^2 \dfrac{r(t-r)}{t^2(t+1)} = (20-10)^2 \dfrac{2(8-2)}{8^2(8+1)} = 2.083 \\ \sigma_X = \sqrt{\sigma_X^2} = \sqrt{2.083} = 1.4434 \\ M = a + (b-a)\dfrac{r-1}{t-2} = 10 + (20-10)\dfrac{2-1}{8-2} = 11.667 \end{array}\right.$$

A. Original PERT Approximations

$$m_X = \frac{a + 4M + b}{6} = \frac{10 + 4 \times 11.667 + 20}{6} = 12.778 \quad (\approx 12.5)$$

$$\sigma_X = \frac{b-a}{6} = \frac{20-10}{6} = 1.67 \quad (\approx 1.44)$$

B. PERRY & GREIG Approximations

$$m_X = \frac{p_5 + 0.95M + p_{95}}{2.95} = \frac{10.53375 + 0.95 \times 11.667 + 15.20703}{2.95} = 12.48 \quad (\approx 12.5)$$

$$\sigma_X = \frac{p_{95} - p_5}{3.25} = \frac{15.20703 - 10.53375}{3.25} = 1.438 \quad (\approx 1.44)$$

This example illustrates that, given perfect 3-point estimates, the Perry and Greig approximations are almost exact, and the PERT approximations are reasonably close.

Properties of the Moments of Random Variables

P.G. Ioannou January 31, 2003

Note: Formulas involving summations apply to *discrete* random variables while formulas with integrals apply to *continuous* random variables. The symbols a and b indicate constants.

Expected Value

(1) $E[X] = m_X = \sum_{\text{all } x} x p_X(x)$

(2) $E[X] = m_X = \int_{-\infty}^{+\infty} x f_X(x) dx$

(3) $E[g(X)] = \sum_{\text{all } x} g(x) p_X(x)$

(4) $E[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$

(5) $E[g(X)] \neq g(E[X])$ (unless $g(X)$ is a linear function of X)

(6) $E[aX] = aE[X]$

(7) $E[b] = b$

(8) $E[aX + b] = aE[X] + b$

(9) $E[g_1(X) + g_2(X)] = E[g_1(X)] + E[g_2(X)]$

(10) $E[X_1 + X_2 + \ldots + X_n] = E[X_1] + E[X_2] + \ldots + E[X_n]$

(11) $E[a_1X_1 + a_2X_2 + \ldots + a_nX_n] = a_1E[X_1] + a_2E[X_2] + \ldots + a_nE[X_n]$

(12)
$$\begin{aligned} E[X] &= \int_{-\infty}^{+\infty} x f_X(x) dx \\ &= \int_{-\infty}^{+\infty} x \{ \int_{-\infty}^{+\infty} f_{XY}(x, y) dy \} dx \\ &= \int_{-\infty}^{+\infty} \{ \int_{-\infty}^{+\infty} x f_{X|Y}(x, y) dx \} f_Y(y) dy \\ &= \int_{-\infty}^{+\infty} E[X|Y] f_Y(y) dy \\ &= E[E[X|Y]] \end{aligned}$$

© 2024 *Photios G. Ioannou*

Variance

(13) $Var[X] = \sigma_X^2 = \sum_{\text{all } x} (x - E[X])^2 p_X(x)$

(14) $Var[X] = \sigma_X^2 = \int_{-\infty}^{+\infty} (x - E[X])^2 f_X(x) dx$

(15) $Var[X] = E[(x - E[X])^2]$

(16) $Var[X] = E[X^2] - (E[X])^2$ (easiest way to calculate $Var[X]$)

(17) $Var[aX] = a^2 Var[X]$

(18) $Var[b] = 0$

(19) $Var[aX + b] = a^2 Var[X]$

Standard Deviation

(20) $\sigma_X = SD[X] = \sqrt{Var[X]} = \sqrt{\sigma_X^2}$

Coefficient of Variation

(21) $V_X = \frac{\sigma_X}{m_X}$

Variance — Covariance — Correlation Coefficient

(22) $Cov[X, Y] = \sigma_{XY} = \sum_{\text{all } x} \sum_{\text{all } y} (x - E[X])(y - E[Y]) p_{XY}(x, y)$

(23) $Cov[X, Y] = \sigma_{XY} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - E[X])(y - E[Y]) f_{XY}(x, y) dx dy$

(24) $Cov[X, Y] = E[(x - E[X])(y - E[Y])]$

(25) $Cov[X, Y] = E[XY] - E[X]E[Y]$ (easiest way to calculate $Cov[X, Y]$)

(26) $\rho_{XY} = \dfrac{Cov[X,Y]}{\sigma_X \sigma_Y} = \dfrac{\sigma_{XY}}{\sigma_X \sigma_Y} \qquad -1 \le \rho_{XY} \le 1$

If X and Y are mutually independent then $\rho_{XY} = 0$

If $(0 < \rho_{XY} \le +1)$ then X and Y are positively correlated

If $(-1 \le \rho_{XY} < 0)$ then X and Y are negatively correlated

(27) $Cov[X,a] = 0$

(28) $Cov[X,X] = \sigma_{XX} = Var[X] = \sigma_X^2 \quad (\rightarrow \rho_{XX} = 1)$

(29) $Cov[aX + b, cY + d] = acCov[X,Y]$

(30) $\rho_{aX+b,cY+d} = \rho_{XY}$

(31) $$\begin{aligned} Var[aX + bY] &= a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_{XY} \\ &= a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\rho_{XY}\sigma_X\sigma_Y \end{aligned}$$

(32) $Var[X \pm Y] = \sigma_X^2 + \sigma_Y^2 \pm 2\sigma_{XY}$

(33) $$\begin{aligned} Var[\sum_{i=1}^{n} X_i] &= \sum_{i=1}^{n}\sum_{j=1}^{n} Cov[X_i, X_j] \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_{X_i X_j} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{X_i X_j}\sigma_{X_i}\sigma_{X_j} \\ &= \sum_{i=1}^{n} Var[X_i] + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i\neq j}}^{n} Cov[X_i, X_j] \\ &= \sum_{i=1}^{n} Var[X_i] + 2\sum_{i=1}^{n}\sum_{j=i+1}^{n} \rho_{X_i X_j}\sigma_{X_i}\sigma_{X_j} \end{aligned}$$

$$
\begin{aligned}
(34)\quad Var[\sum_{i=1}^{n} a_i X_i] &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j Cov[X_i, X_j] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \sigma_{X_i X_j} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \rho_{X_i X_j} \sigma_{X_i} \sigma_{X_j} \\
&= \sum_{i=1}^{n} a_i^2 Var[X_i] + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i \neq j}}^{n} a_i a_j Cov[X_i, X_j] \\
&= \sum_{i=1}^{n} a_i^2 Var[X_i] + 2\sum_{i=1}^{n}\sum_{j=i+1}^{n} a_i a_j \rho_{X_i X_j} \sigma_{X_i} \sigma_{X_j}
\end{aligned}
$$

The following two formulas apply if the random variables $X_1, X_2, \ldots, X_n$ are independent:

$$(35)\quad Var[X_1 + X_2 + \ldots + X_n] = Var[X_1] + Var[X_2] + \ldots + Var[X_n]$$

$$(36)\quad Var[a_1 X_1 + a_2 X_2 + \ldots + a_n X_n] = a_1^2 Var[X_1] + a_2^2 Var[X_2] + \ldots + a_n^2 Var[X_n]$$

Moments of X, Y Functions

$$(37)\quad E[g(X,Y)] = \sum_{\text{all } x}\sum_{\text{all } y} g(X,Y)\, p_{XY}(x,y)$$

$$(38)\quad E[g(X,Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(X,Y)\, f_{XY}(x,y)\, dx dy$$

$$(39)\quad E[g(X,Y)] \neq g(E[X], E[Y]) \qquad \text{(unless } g(X,Y) \text{ is a linear function of } X, Y)$$

Moments of a Product

$$(40)\quad E[XY] = E[X]E[Y] + Cov[X,Y]$$

$$(41)\quad E[XY] = E[X]E[Y] \qquad \text{(where } X \text{ \& } Y \text{ are independent)}$$

$$(42)\quad E[\frac{X}{Y}] \neq \frac{E[X]}{E[Y]} \qquad \text{(even if } X \text{ \& } Y \text{ are independent)}$$

$$(43)\quad Var[XY] = m_X^2 \sigma_Y^2 + m_Y^2 \sigma_X^2 + \sigma_X^2 \sigma_Y^2 \qquad \text{(where } X \text{ \& } Y \text{ are independent)}$$

$$(44)\quad V_Z^2 = V_X^2 + V_Y^2 + V_X^2 V_Y^2 \qquad \text{(where } Z = XY, \text{ and } X \text{ \& } Y \text{ are independent)}$$

Use of Expected Values in Games

We will later see how expected value is the criterion for selecting among alternative construction decisions. The use and usefulness of this criterion is interestingly illustrated in the evaluation of games of chance.

Flipping Coins for Pennies

For a wager of one cent, one receives two cents if one guesses correctly whether two flipped coins show the same (even) or different (odd) faces.

Upon calling "even", there are two events with two payoffs similar to a Bernoulli random variable, where *even* is equivalent to the value 1 (*success*) and *odd* is equivalent to the value 0 (*failure*). Obviously, the probability of success, $p = 0.5$.

Let X be the underlying Bernoulli random variable and R be the return of the wager in pennies. The expected return is,

$$E[R] = E[2X] = \sum_{all\ i} 2x_i P[X = x_i] = 2\sum_{all\ i} x_i P[X = x_i] = 2\{1p + 0(1-p)\} = 2p = 2\frac{1}{2} = 1\ cent$$

For a wager of 1 cent, the expected return is also 1 cent. Therefore, this is a fair game.

Las Vegas Roulette

On a Las Vegas roulette table, there are 38 numbers (1 through 36 plus 0 and 00). You receive \$36 for each \$1 bet if the ball lands on the number on which you bet.

This game can also be described by using a Bernoulli random variable, where *success* is the event where the ball lands on the number chosen, and *failure* is the event where the ball lands anywhere else. Clearly, in this case the probability of success is $p = 1/38$.

Let X be the underlying Bernoulli random variable and R be the return of the bet in dollars. The expected return is,

$$E[R] = E[36X] = 36\{1p + 0(1-p)\} = 36p = 36\frac{1}{38} = \$0.947$$

On each dollar bet, the return is \$0.947 and the "house" expects \$0.053 or 5.3% as a cut. This is not a fair game.

Game of Twenty-One

One plays cards against a dealer of the gambling house. You and the dealer are each dealt two cards. You can see both of yours and one of the dealers. You can ask the dealer to deal you additional cards, one at a time, as long as the sum of all the cards you have been dealt (including the original two) does not exceed 21. If you go over 21, you lose and the game is over. If you *stay* (i.e., the sum of all your cards is 21 or less and you refuse additional cards), then the dealer

must draw additional cards if the sum of all his cards is 16 or less. If during this process the dealer's sum becomes any number from 17 to 21 the dealer must *stay*. If he goes over 21 then he loses. You win if the sum of your cards is greater than the dealers but not over 21.

Question: Should you draw a card is you have a sum of 16 and the dealer also has a sum of 16?

Find your expected value of receiving a single additional card. For both you and the dealer:

Let *Y* represent the total sum of your cards assuming that you draw an additional card and let Z represent the dealer's sum after he draws a card (notice that the dealer must stay at that point). The dealer's total, *Z*, is independent of yours, *Y*, if you are not above 21. You will win under the following circumstances:

y	$P[Y=y]$	Win if Z is	$P[win \mid Y=y]$	$P[win \mid Y=y]$ $\cdot P[Y=y]$	$P[tie \mid Y=y]$	$P[tie \mid Y=y]$ $\cdot P[Y=y]$
17	1/13	>21	8/13	8/169	1/13	1/169
18	1/13	17,>21	9/13	9/169	1/13	1/169
19	1/13	17,18,>21	10/13	10/169	1/13	1/169
20	1/13	17,18,19,>21	11/13	11/169	1/13	1/169
21	1/13	17,18,19,20,>21	12/13	12/169	1/13	1/169
>21	8/13	none	0	0	0	0
				$\Sigma = 50/169$ $= 0.296$		$\Sigma = 5/169$ $= 0.030$

$$P[\text{win}] = \sum_i P[\text{win} \mid Y=y]P[Y=y] = 50/169 = 0.296$$

$$P[\text{tie}] = \sum_i P[\text{tie} \mid Y=y]P[Y=y] = 5/169 = 0.030$$

For \$1 bet, $E[R] = \$2(0.296) + \$1(0.030) = 0.592 + 0.030 = \$0.622$

If you draw no card, you win if the dealer goes over 21,

$$P[Z > 21] = 8/13 = 0.615$$

For \$1 bet,

$$E[R] = \$2(0.615) = \$1.231 > \$0.622$$

Conclusion: Do not draw card.

(Caution: The situation described here is different from real life because at a casino you can only see the dealer's top card. The dealer's hidden card is called his *hole* card. A player's best choice upon seeing a top card of 6 and not knowing the hole card is more complex.)

Example—How to Calculate Covariance

$p_{X,Y}(x,y)$	Y			
X	0	1	2	$p_X(x)$
0	2/9	1/18	1/18	1/3
1	1/18	2/9	1/18	1/3
2	1/18	1/18	2/9	1/3
$p_Y(y)$	1/3	1/3	1/3	1.0

$$E[X] = E[Y] = 0\times\frac{1}{3} + 1\times\frac{1}{3} + 2\times\frac{1}{3} = 1$$

$$E[X^2] = E[Y^2] = 0^2\times\frac{1}{3} + 1^2\times\frac{1}{3} + 2^2\times\frac{1}{3} = \frac{5}{3}$$

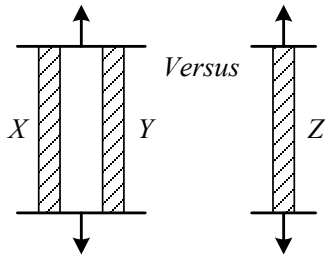
$$Var[X] = Var[Y] = \frac{5}{3} - 1^2 = \frac{2}{3}$$

$$\begin{aligned} E[XY] = \; & 0\times 0\times\frac{2}{9} \quad + 0\times 1\times\frac{1}{18} \quad + 0\times 2\times\frac{1}{18} \\ & + 1\times 0\times\frac{1}{18} \quad + 1\times 1\times\frac{2}{9} \quad + 1\times 2\times\frac{1}{18} \\ & + 2\times 0\times\frac{1}{18} \quad + 2\times 1\times\frac{1}{18} \quad + 2\times 2\times\frac{2}{9} \\ = \; & \frac{2}{9} + \frac{1}{9} + \frac{1}{9} + \frac{8}{9} = \frac{4}{3} \end{aligned}$$

$$Cov[X,Y] = \frac{4}{3} - 1\times 1 = \frac{1}{3} \quad \Rightarrow \quad \rho_{XY} = \frac{\frac{1}{3}}{\sqrt{\frac{2}{3}}\sqrt{\frac{2}{3}}} = \frac{1}{2}$$

Notice that because X and Y are not independent, $E[XY] = \frac{4}{3} \neq E[X]E[Y] = 1$

Example—Effect of Covariance/Correlation



X, Y, Z = yield strength of steel rods

Given:

$$m_X = m_Y = 30 \cdot 10^6 lbs$$
$$m_Z = 60 \cdot 10^6 lbs$$

Rods are made of same material, hence:

$$V_X = V_Y = V_Z = \frac{\sigma}{m} = 0.06667$$

The variability in steel rod strength can be determined as

$$\sigma_Z = V_Z m_Z = 0.06667 \times 60 \cdot 10^6 lbs$$
$$\sigma_Z^2 = 16 \cdot 10^{12} (lbs)^2$$

$$\sigma_X = \sigma_Y = V_X m_X = V_Y m_Y = 0.06667 \times 30 \cdot 10^6 = 2 \cdot 10^6 lbs$$
$$\sigma_X^2 = \sigma_Y^2 = 4 \cdot 10^{12} (lbs)^2$$

Let W be the combined strength of two rods, $W = X + Y$

$$m_W = m_X + m_Y = 60 \cdot 10^6 lbs \quad \Rightarrow m_W = m_Z$$

$$\sigma_W^2 = Var[W] = Var[X + Y] = \sigma_X^2 + \sigma_Y^2 + 2\rho_{XY}\sigma_X\sigma_Y$$
$$= 4 \cdot 10^{12} + 4 \cdot 10^{12} + 2 \times 2 \cdot 10^6 \times 2 \cdot 10^6 \times \rho_{XY} = 8 \cdot 10^{12} + 8 \cdot 10^{12} \rho_{XY}$$
$$= 8 \cdot 10^{12} (1 + \rho_{XY})$$

If X and Y are independent,

$$\rho_{XY} = 0 \quad \Rightarrow \quad \sigma_W^2 = 8 \cdot 10^{12} = \frac{\sigma_Z^2}{2} \quad (!)$$

In any event:

$$\sigma_W^2 = 8 \cdot 10^{12} (1 + \rho_{X,Y}) \quad \leq \quad \sigma_Z^2 = 16 \cdot 10^2$$

Thus, the use of smaller diameter reinforcing bars should be preferred as it increases reliability.

Steps in Making a Decision under Uncertainty

1. List alternative actions.

2. List the possible outcomes of the alternative actions which will influence the value of the actions.

3. Describe these outcomes as the events which make up the sample points in the sample space of each alternative action. These may be compound events made up of joint occurrences of simple events.

4. Determine the probability, $P(x_i)$, of each sample point. This requires the input of probabilities of simple events and conditional probabilities of compound events.

5. Determine the value, x_i, of each sample point for each alternative action.

6. Calculate the expected value of each alternative action by

$$E[X] = \sum_{\text{all } i} x_i P(x_i)$$

7. Select the alternative action which has the highest expected value.

Expected Value in Construction Decisions

The risks associated with construction operations can be equal to those in games of chance.

As in resolving the odds and expectations in games of chance, expected value is the criteria for evaluating construction decisions.

Example An owner is told that the unit cost for the construction of a new facility, *C*, would have the following outcomes and probabilities.

c (\$/SF)	$P[C = c]$	$cP[C = c]$
45	0.2	9
50	0.3	15
55	0.4	22
60	0.1	6
		$\Sigma = 52$

The expected unit cost of construction would be

$$E[C] = \sum_{\text{all } c} cP[C = c] = 45(0.2) + 50(0.3) + 55(0.4) + 60(0.1) = \$52 / SF$$

Example—Probabilistic Expense Allocation for Tractor

A contractor purchases a tractor for use in a quarry at an initial cost of \$35,000. The lifetime L over which the tractor will be productive is uncertain and follows the probability mass function given below. The operating cost H (which includes maintenance and fuel costs) is related to the tractor's productive life by the following table of probabilities. The expected salvage value at the end of the tractor's productive life is \$10,000. Insurance and taxes are \$2,000 per year.

$P[H=h \mid L=l]$	Operating Costs, h (\$/hr)		
Lifetime, l (hrs)	4.00	5.00	6.00
4000	-	-	1.00
6000	0.1	0.3	0.6
8000	0.3	0.6	0.1
10000	0.7	0.3	-

$P[L=l]$
0.1
0.2
0.4
0.3

What hourly costs should the contractor assign to the tractor? Assume 2,000 hr/yr operation.

In standard practice, all costs are allocated as hourly costs of equipment use. The total hourly cost, C, is the sum of the hourly ownership cost Q and hourly operating costs H:

$$C = Q + H \quad \Rightarrow E[C] = E[Q] + E[H]$$

The ownership cost equals the tractor's initial cost minus its salvage value plus the annual interest and taxes. These are allocated per hour of operating time. Thus, the hourly ownership cost Q (\$/hr) is a function of life L,

$$Q = g(L) = \frac{\text{initial cost - salvage value}}{L \text{ hours of life}} + \frac{\text{annual insurance \& taxes}}{2{,}000\ hr/yr}$$

$$= \frac{35{,}000 - 10{,}000}{L} + \frac{2{,}000}{2{,}000}$$

$$= \frac{25{,}000}{L} + 1 \quad (\$/hr)$$

$$E[Q] = E[g(L)] = \sum_{\text{all } l} g(l)P[L=l] = \sum_{\text{all } l} \left(\frac{25{,}000}{l} + 1\right)P[L=l]$$

$$= \left(\frac{25{,}000}{4{,}000} + 1\right)0.1 + \left(\frac{25{,}000}{6{,}000} + 1\right)0.2 + \left(\frac{25{,}000}{8{,}000} + 1\right)0.4 + \left(\frac{25{,}000}{10{,}000} + 1\right)0.3$$

$$= \$4.458/hr$$

There are two ways to calculate the expected value of the hourly operating costs, H. The first is straightforward and requires that we calculate the marginal probability mass function of H.

The second, requires that we calculate the conditional expected values $E[H \mid L=l]$ and then use these and the given marginal PMF $P[L=l]$, to calculate the unconditional $E[H]$,

$$E[H] = \sum_{all\ \ l} E[H \mid L=l]P[L=l] = \sum_{all\ \ l} \left\{ \sum_{\text{all } h} hP[H=h \mid L=l] \right\} P[L=l]$$

To use the first method, we use the given probabilities to calculate the joint probability mass function for H and L, and from that the marginal probability mass function of H, as follows:

$P[H=h \mid L=l]$	Operating Costs, h (\$/hr)		
Lifetime, l (hrs)	4.00	5.00	6.00
4000	-	-	1.00
6000	0.1	0.3	0.6
8000	0.3	0.6	0.1
10000	0.7	0.3	-

$P[L=l]$
0.1
0.2
0.4
0.3

$P[H=h, L=l]$	Operating Costs, h (\$/hr)			
Lifetime, l (hrs)	4.00	5.00	6.00	$P[L=l]$
4000	-	-	0.10	0.1
6000	0.02	0.06	0.12	0.2
8000	0.12	0.24	0.04	0.4
10000	0.21	0.09	-	0.3
$P[H=h]$	0.35	0.39	0.26	1.0

$$E[H]=\sum_{\text{all } r} hP[H=h]=4\times 0.35+5\times 0.39+6\times 0.26=\$4.91/hr$$

To use the second method, we calculate the conditional expected values first,

$$E[H \mid L=l]=\sum_{\text{all } r} hP[H=h \mid L=l] \quad \Rightarrow$$

$$E[H \mid L=4{,}000] = 4\times 0.0+5\times 0.0+6\times 1.0=\$6.0/hr$$

$$E[H \mid L=6{,}000] = 4\times 0.1+5\times 0.3+6\times 0.6=\$5.5/hr$$

$$E[H \mid L=8{,}000] = 4\times 0.3+5\times 0.6+6\times 0.1=\$4.8/hr$$

$$E[H \mid L=10{,}000] = 4\times 0.7+5\times 0.3+6\times 0.0=\$4.3/hr$$

Then, we use the conditional expected values $E[H \mid L=l]$ to calculate the unconditional expected value of H,

$$E[H]=\sum_{all\ \ l} E[H \mid L=l]P[L=l]$$

$$=6\times 0.1+5.5\times 0.2+4.8\times 0.4+4.3\times 0.3=\$4.91/hr$$

The expected total hourly cost that the contractor should assign to the tractor, $E[C]$, is the sum of the expected hourly ownership cost $E[Q]$ and expected operating costs $E[H]$,

$$E[C]=E[Q]+E[H]=4.458+4.910=\$9.368/hr$$

Statistics: Sample Mean and Variance

It is quite common that when a random variable is sampled, the *sample mean* or *sample average* can be calculated as an *approximation* of the true mean of the random variable. The *sample mean* or *sample average* is defined as follows:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Example. We are interested in the mean 28-day cylinder strength of a particular concrete mix design. One can estimate that mean by mixing 10 batches, filling 3 cylinders from each, and testing them after curing. Their average is an estimate of the true mean strength of the batch design. Obviously, repeating the experiment will produce a sample mean which would be expected to differ slightly from the first or any subsequent experiments. Statistical methods for designing experiments to produce estimates of means to a certain accuracy will be described later.

As with the sample mean, we often use experimental outcomes to compute the *sample variance* S^2 as an approximation of the variance of the random variable. To maintain the analogy, the square root of the sample variance is called *sample standard deviation*, *S*, and is an estimator of the standard deviation of the random variable.

There are two formulas for the *sample variance* S^2 used in practice, depending on whether we divide the sum of the squares of the sample deviations from the average by *n* or *n*-1.

The sample variance S_n^2 uses *n* as the divisor:

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - (\bar{X})^2 = \overline{X^2} - \bar{X}^2$$

As shown later, although S_n^2 is an *efficient* estimator of σ^2, it is *biased* and underestimates σ^2

$$E[S_n^2] = \frac{n-1}{n}\sigma^2$$

As a result, we define the unbiased estimator S_{n-1}^2,

$$S_{n-1}^2 = \frac{n}{n-1}S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n-1}\sum_{i=1}^{n} X_i^2 - \frac{n}{n-1}(\bar{X})^2 = \frac{n}{n-1}(\overline{X^2} - \bar{X}^2)$$

Thus, by definition, S_{n-1}^2 is unbiased, i.e., $E[S_{n-1}^2] = \sigma^2$. However, it is less efficient than S_n^2,

$$Var[S_{n-1}^2] > Var[S_n^2]$$

Consequently, the choice between S_n^2 and S_{n-1}^2 is not clear cut. S_n^2 has less variability but on average tends to underestimate σ^2. Conversely, S_{n-1}^2 tends to estimate σ^2 correctly but has more variability. In practice, we usually prefer to use the unbiased estimator S_{n-1}^2.

Example. We wish to estimate the average cycle-time for a front-end loader. We measure with a stopwatch a total of 20 cycles to get the 20 samples below.

x_i (sec)	x_i^2 (sec^2)
18.6	345.96
16.3	265.69
17.0	289.00
17.2	295.84
18.3	334.89
18.9	357.21
15.8	249.64
16.8	282.24
16.9	285.61
17.2	295.84
17.3	299.29
15.9	252.81
16.4	268.96
17.6	309.76
20.3	412.09
17.4	302.76
18.5	342.25
17.9	320.41
18.6	345.96
18.0	324.00
$\sum_{i=1}^{n} x_i = 350.9$	$\sum_{i=1}^{n} x_i^2 = 6{,}180.21$

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{20}350.9 = 17.545\,\text{sec}$$

$$\overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 = \frac{1}{20}6{,}180.21 = 309.0105\,\text{sec}^2$$

$$s_n^2 = \overline{x^2} - \overline{x}^2 = 309.0105 - 17.545^2 = 1.1835\,\text{sec}^2$$

$$s_{n-1}^2 = \frac{n}{n-1}s_n^2 = \frac{20}{19}1.1835 = 1.2458\,\text{sec}^2$$

$$s_{n-1} = \sqrt{s_{n-1}^2} = \sqrt{1.2458} = 1.1161\,\text{sec}$$

We can estimate the cycle time as $\overline{x} \pm s = 17.55 \pm 1.12\,\text{sec}$.

The *relative variability* of the data can be measured by a dimensionless statistic called the *sample coefficient of variation*, which is the ratio of the sample standard deviation to the sample mean. For the front-end loader data, we have,

$$v = \frac{s}{\bar{x}} = \frac{1.1161}{17.545} = 6.36\%$$

Data Reduction for Random Variables

Samples of random variables serve as inputs to the understanding of the random variable's probability distributions. There are many different forms of description for a random variable and different statistical methods which will reduce the raw data samples to these useful forms of description. The samples can be sorted, or ordered, and shown graphically as an approximation of the *PMF*, *PDF*, and *CDF* of the true probability distribution.

Graphical Displays

Histogram. One of the simplest ways of describing a random variable through its raw data is the histogram, a type of bar chart. A histogram shows the number of samples that fall within each of a set of intervals. A histogram is obtained by (1) dividing the range of samples into a number of intervals, (2) counting the samples within each interval, and (3) plotting the result.

To illustrate the process, consider the following example. *Acorn* is a construction contractor whose bidding experience over two years is summarized by the data for 48 projects in the table below. For each project, this table shows Acorn's estimated cost, Acorn's bid, and the lowest bid submitted by Acorn's competitors. The variable of interest in this example is the *LBC* for each project, i.e., the "low-bid-to-cost ratio", which for each past project is defined as:

LBC = (competitors' low bid) / (Acorn's estimated cost).

At first, we examine the data and establish their range: from 0.7723 to 1.2391.

Then we need to subdivide this range into a number of intervals, *k*. Obviously, the number of intervals *k* affects the width of each interval. For a given data range, the larger the number of intervals *k*, the narrower each interval must be. Consequently, as *k* increases, the number of samples that fall within each interval decreases, and the histogram becomes less smooth and more irregular. Thus, when choosing *k* we must strike a balance because larger values that increase accuracy, or resolution, and smaller values that produce smoother histograms that resemble the underlying probability distribution more closely.

Because the choice of interval can greatly influence the appearance of a histogram and the information it gives the viewer, it is often best to use a number of intervals *k* close to that proposed by Sturges (Benjamin and Cornell, p. 8):

$$k = 1 + 3.3 \log_{10} n$$

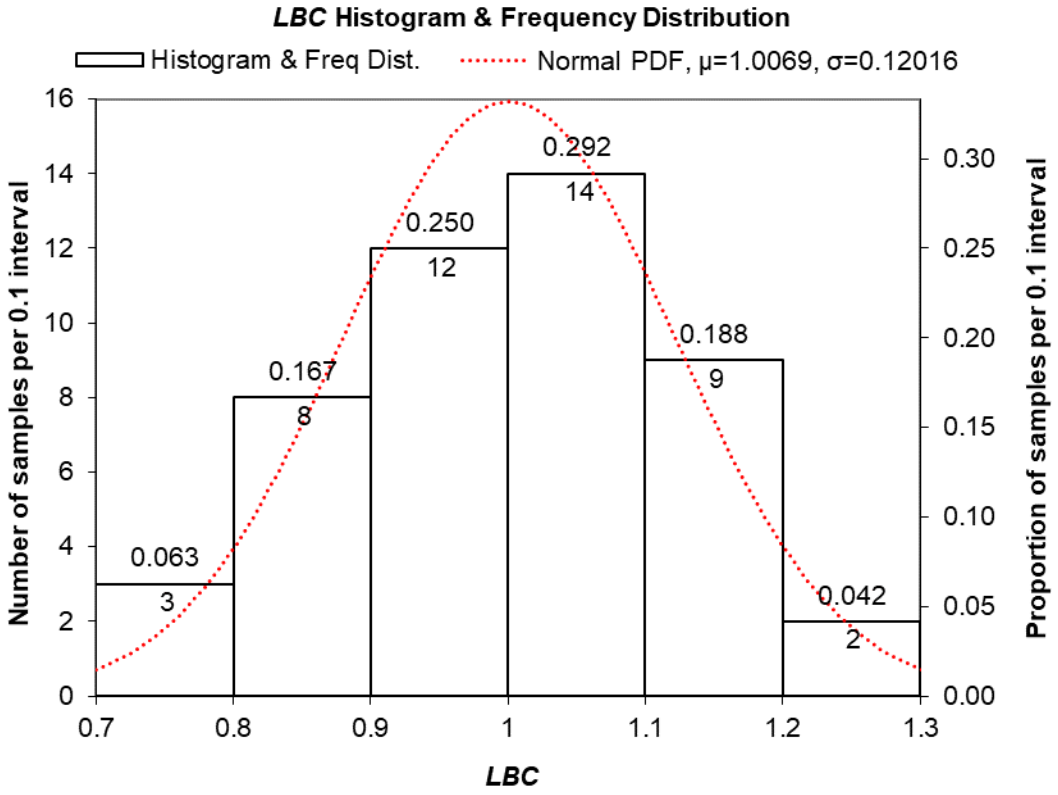
where *n* is the number of samples available. In this case, $k = 1+3.3 \log(48) = 6.55$.

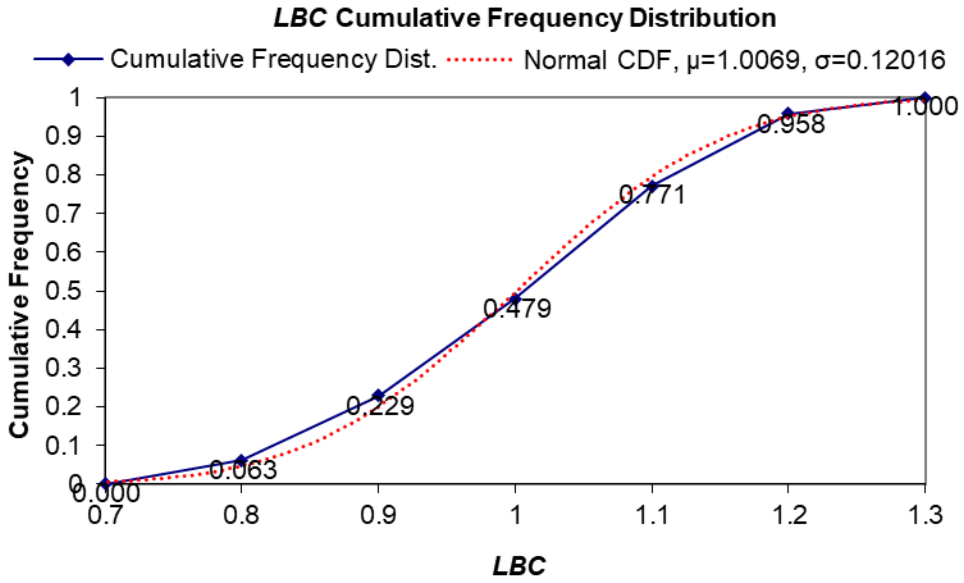
Let us use 6 intervals, each with a width of 0.10, ranging from 0.7 to 1.3. To construct a histogram, we count the number of samples in each interval and we plot the results in a column chart. In the figures below, the left ordinate shows the number of samples (histogram) and the right ordinate their proportion to the total number of samples (frequency).

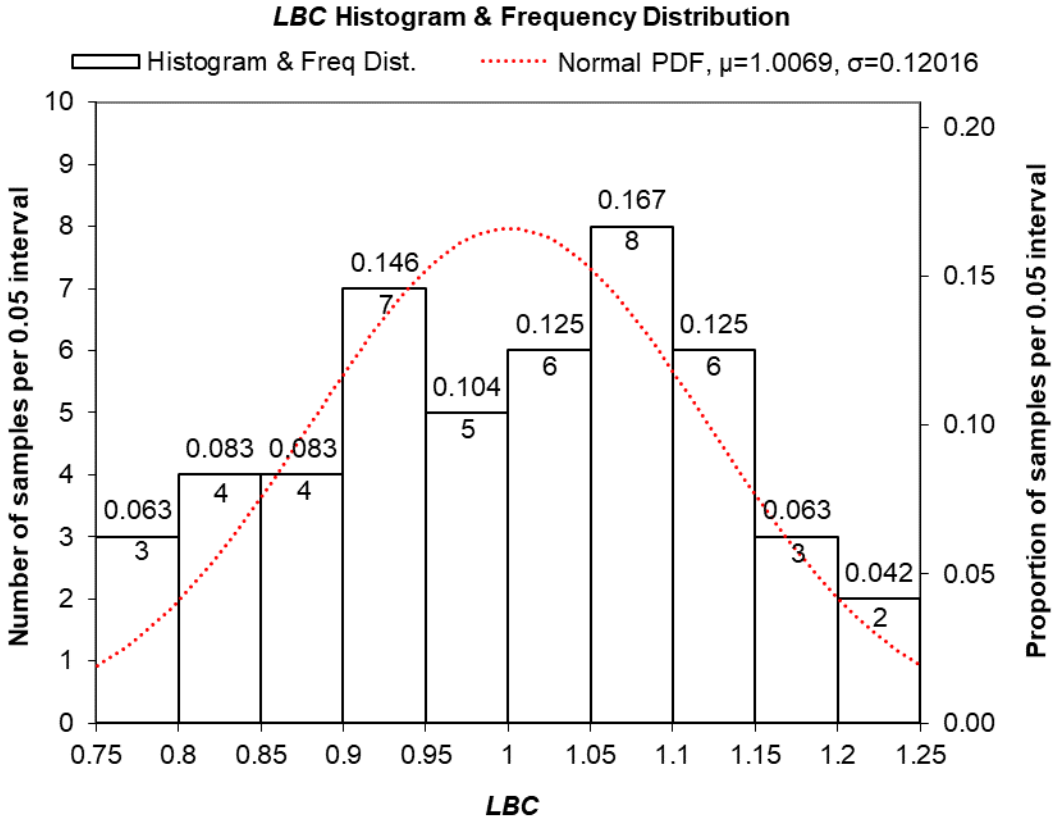
It may sometimes be advantageous to make histograms with different intervals. A larger number of intervals, of size 0.05, result in the figures that follow and give a slightly different impression of *LBC*.

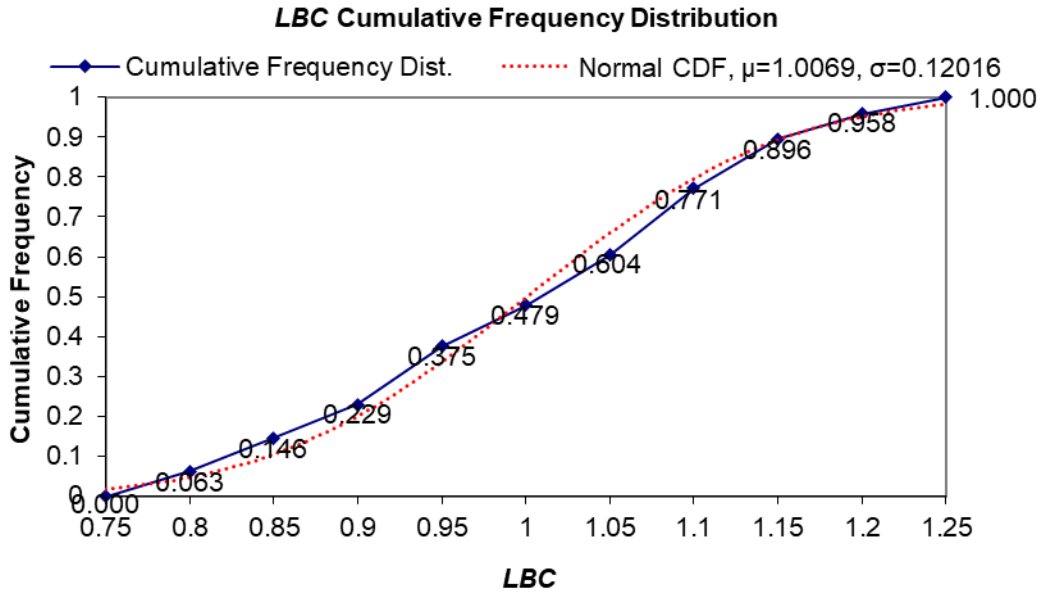
Acorn's Bidding History

Job No	Acorn's Cost	Acorn's Bid	Competitors' Low Bid	LBC
1	45,943	53,000	50,950	1.1089
2	6,343	7,450	6,655	1.0491
3	6,586	8,220	6,276	0.9529
4	6,757	8,450	7,400	1.0951
5	374,918	396,500	416,000	1.1095
6	136,775	149,650	148,600	1.0864
7	5,064	5,950	4,500	0.8886
8	87,535	94,500	96,000	1.0967
9	12,303	13,350	15,000	1.2192
10	50,884	58,500	59,800	1.1752
11	6,588	7,650	7,015	1.0648
12	56,368	65,000	46,900	0.8320
13	31,750	35,825	33,660	1.0601
14	46,020	53,000	53,600	1.1647
15	193,478	225,000	239,742	1.2391
16	515,613	567,500	579,600	1.1240
17	189,577	219,500	213,400	1.1256
18	230,162	260,000	193,800	0.8420
19	195,398	216,000	194,000	0.9928
20	22,275	26,750	22,844	1.0255
21	24,383	28,000	20,000	0.8202
22	5,073	5,500	5,400	1.0644
23	56,442	62,750	44,800	0.7937
24	42,411	50,975	41,600	0.9808
25	42,065	47,500	44,300	1.0531
26	359,068	402,000	360,000	1.0025
27	8,155	9,850	9,069	1.1120
28	39,467	45,250	33,700	0.8538
29	263,552	287,500	239,650	0.9093
30	81,826	108,750	63,200	0.7723
31	8,994	10,750	8,450	0.9395
32	13,985	16,000	11,789	0.8429
33	262,185	310,000	268,876	1.0255
34	22,414	24,950	22,000	0.9815
35	21,274	23,350	22,231	1.0449
36	434,925	480,000	434,300	0.9985
37	482,774	511,750	532,759	1.1035
38	163,268	179,750	129,342	0.7922
39	35,179	40,590	41,700	1.1853
40	92,130	99,750	80,200	0.8705
41	85,998	94,475	81,000	0.9418
42	266,274	288,000	269,967	1.0138
43	16,416	17,950	14,741	0.8979
44	686,409	741,000	623,600	0.9084
45	266,781	289,000	249,240	0.9342
46	281,405	298,000	264,872	0.9412
47	50,906	55,555	54,964	1.0797
48	459,391	491,000	421,400	0.9173
			E[LBC] ~=	1.00069
			SD[LBC] ~=	0.12016


LBC Histogram & Frequency Distribution
Histogram & Freq Dist.
Normal PDF, μ=1.0069, σ=0.12016
Number of samples per 0.1 interval
Proportion of samples per 0.1 interval
0.063
3
0.167
8
0.250
12
0.292
14
0.188
9
0.042
2
LBC


LBC Cumulative Frequency Distribution
Cumulative Frequency Dist.
Normal CDF, μ=1.0069, σ=0.12016
Cumulative Frequency
0.000
0.063
0.229
0.479
0.771
0.958
1.000
LBC


LBC Histogram & Frequency Distribution
Histogram & Freq Dist.
Normal PDF, μ=1.0069, σ=0.12016
Number of samples per 0.05 interval
Proportion of samples per 0.05 interval
LBC
0.063
3
0.083
4
0.083
4
0.146
7
0.104
5
0.125
6
0.167
8
0.125
6
0.063
3
0.042
2


LBC Cumulative Frequency Distribution
Cumulative Frequency Dist.
Normal CDF, μ=1.0069, σ=0.12016
Cumulative Frequency
LBC
0.000
0.063
0.146
0.229
0.375
0.479
0.604
0.771
0.896
0.958
1.000

Frequency Distribution

A *frequency distribution* graph shows the proportion of the total number of samples that fall within each interval. Thus, it is identical to the corresponding histogram, except that the ordinate values are divided by the total number of samples. For the *LBC* example, the frequency distribution shows that the proportion in the interval 0.7 to 0.8 is 3/48 = 0.063. In the figures above, the scale for the frequency distribution is shown on the right ordinate. Notice that the sum of all ordinate values for the frequency distribution graph is 48/48 = 1. Thus, a frequency distribution graph is similar to a probability mass function where the sum of all masses equals 1. This is different from a probability density function, because the area under the frequency distribution graph does not equal 1, but rather the width of one histogram interval.

Cumulative Frequency Distribution based on a Histogram

Shown below each histogram/frequency distribution graph is the corresponding *cumulative frequency distribution* graph for *LBC*. For each of the values x that form the upper boundaries of the histogram intervals, this graph shows the relative proportion of the sample points that were less than or equal to x. Thus, the cumulative frequency distribution graph is formed by summing the values shown on the frequency distribution graph.

For the *LBC* example using six 0.1 histogram intervals,

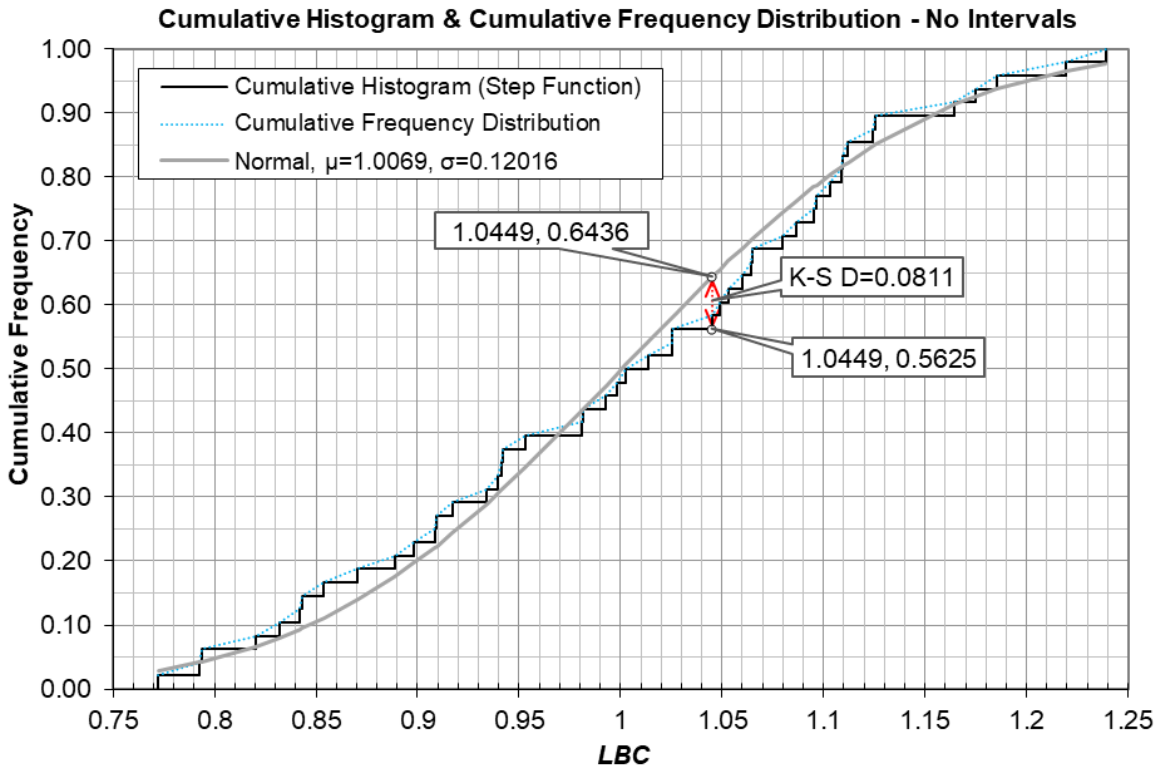
$$F(0.80) = P(X \le 0.80) = 0.000 + 0.063 = 0.063$$
$$F(0.90) = P(X \le 0.90) = 0.063 + 0.167 = 0.230$$
$$F(1.00) = P(X \le 1.00) = 0.230 + 0.250 = 0.480$$
$$F(1.10) = P(X \le 1.10) = 0.480 + 0.292 = 0.772$$
$$F(1.20) = P(X \le 1.20) = 0.772 + 0.188 = 0.960$$
$$F(1.30) = P(X \le 1.30) = 0.960 + 0.042 = 1.000$$

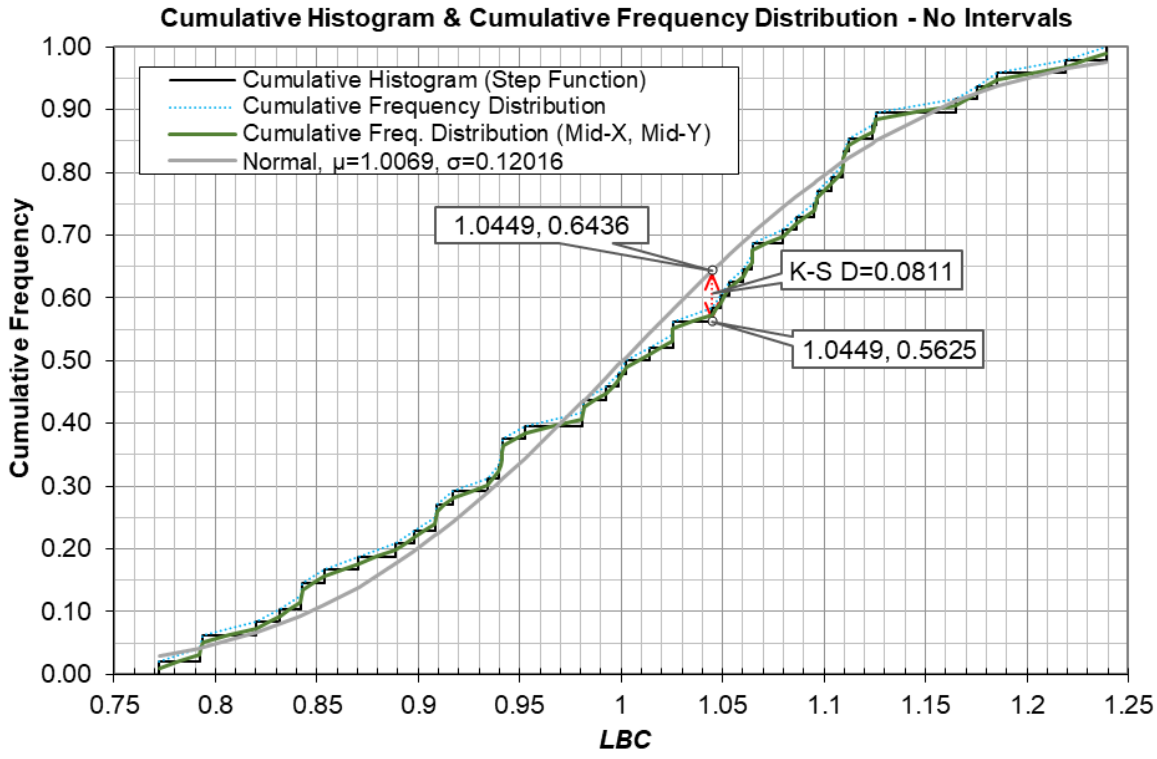
To complete the cumulative frequency distribution graph, we connect the points above with linear segments. This implies that the underlying random variable is continuous and that there could be other points between the *LBC* values shown on the horizontal axis. If the underlying random variable is discrete, and the sample set contained only the values shown on the horizontal axis of histogram, then the cumulative frequency distribution graph would be a step function.

Cumulative Histogram and Cumulative Frequency Distribution based only on Data

A cumulative frequency distribution can also be constructed *without* having to select a number of intervals or bins, k. In this case, we sort all 48 data from the lowest to the highest value and construct a *step function* where the ordinate increases (*jumps up*) by 1/48 at each one of the 48 data values. This step function is called a cumulative histogram and it is the one required for the statistic, D, of the Kolmogorov-Smirnov goodness-of-fit test. Connecting the top-left corners of the step function with linear segments produces a cumulative frequency distribution. Connecting the mid-points of the "steps" and "rises" gives a different cumulative frequency distribution (shifted to the right). Notice that in all these cases, the horizontal width of each *step* is not constant and depends on how close to each other the successive data points happen to be.

© 2024 *Photios G. Ioannou*


Cumulative Histogram & Cumulative Frequency Distribution - No Intervals
Cumulative Histogram (Step Function)
Cumulative Frequency Distribution
Normal, μ=1.0069, σ=0.12016
1.0449, 0.6436
K-S D=0.0811
1.0449, 0.5625
Cumulative Frequency
LBC
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0.75
0.8
0.85
0.9
0.95
1
1.05
1.1
1.15
1.2
1.25


Cumulative Histogram & Cumulative Frequency Distribution - No Intervals
Cumulative Histogram (Step Function)
Cumulative Frequency Distribution
Cumulative Freq. Distribution (Mid-X, Mid-Y)
Normal, μ=1.0069, σ=0.12016
1.0449, 0.6436
K-S D=0.0811
1.0449, 0.5625
Cumulative Frequency
LBC
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0.75
0.8
0.85
0.9
0.95
1
1.05
1.1
1.15
1.2
1.25

Kolmogorov-Smirnov Test

This test is used to decide if a data sample comes from a hypothesized continuous distribution. It is based on the *cumulative histogram* which is a "stair-step function" of the data like the one shown above. In the program EasyFit, the *cumulative histogram* is called the "*empirical cumulative distribution function*" (ECDF).

Assume that we have a random sample x_1, ..., x_n from some distribution with CDF, $F(x)$, and that the sample data have been sorted in ascending order. The empirical CDF is denoted by:

$$F_n(x) = \frac{1}{n}[\text{Number of observations} \le x]$$

Definition of the Kolmogorov-Smirnov (K-S) Statistic

The Kolmogorov-Smirnov (K-S) statistic, D, is based on the *largest vertical difference* between the *theoretical* and the *empirical* cumulative distribution function:

$$D = \max_{1 \le i \le n} \left(\left| F(x_i) - \frac{i-1}{n} \right|, \left| \frac{i}{n} - F(x_i) \right| \right)$$

Hypothesis Testing

The null and the alternative hypotheses are:

- H_0: the data follow the specified distribution;
- H_1: the data do not follow the specified distribution.
  Operating Rule: Reject the null hypothesis, H_0, in favor of H_1, if $D > c$.

The value, c, is called the *critical value* and corresponds to some chosen *significance level*, α. The *level of significance*, α, is the probability of making a mistake and rejecting the null hypothesis, when it is actually true. Thus, smaller values of α, lead to smaller probabilities of making a *Type I* error and rejecting H_0, when H_0 is actually true. However, smaller values of α, lead to larger probabilities of making a *Type II* error and accepting H_0, when H_0 is actually false.

For a chosen *significance level*, α, (typically, 1%, 5%, 10%), the *critical value*, c, depends on the sample size, n, and is obtained from special tables such as the one shown below.

Example

Consider the *LBC* data shown above and the following null hypothesis:

- H_0: The LBC data come from a Normal distribution with μ=1.0069, and σ=0.12016.

In this case, the Kolmogorov-Smirnov (K-S) statistic is: D = 0.0811 (as shown in figure).

For n=48 and α=10%, we obtain the critical value, $c = 1.22 / \sqrt{n} = 0.176$.

Since D = 0.0811 < c = 0.176, we must accept (or, more formally, "we cannot reject") H_0, that the LBC data come from the Normal distribution with μ=1.0069, and σ=0.12016.

Note there are many distributions that cannot be rejected as having produced a given sample. For example, the null hypothesis that the given *LBC* data came from a Gamma distribution can also not be rejected at the level of significance, α=10%.

The tables of critical values for the K-S test are valid only when testing "*whether a data set comes from a completely specified distribution*". If one or more distribution parameters (such as μ=1.0069, and σ=0.12016 for the example above) are estimated from the same data, then the true level of significance, α_{true}, is actually smaller than the one used for the table, $\alpha_{true} < \alpha$, and the probability that the fit (i.e., H_0) will be rejected, when it is actually true, will be lower.

P-Value

The *P*-value equals the level of significance, α, for which $D = c$. Thus, the *P*-value is the level of significance that is the threshold between accepting and rejecting the null hypothesis, H_0.

- H_0 will be accepted for all values of α less than the *P*-value.
- H_0 will be rejected for all values of α greater than the *P*-value.

For the above example of *LBC* data, the *P*-value (reported by *EasyFit*) is very high:

P-value = 88.547%.

This means that the null hypothesis cannot be rejected at all significance levels, $\alpha < 88\%$. To reject the null hypothesis, the level of significance, α, would have to be uncommonly high (above 88%). Thus, this null hypothesis cannot be rejected under normal practices.

Critical Statistic Values for the Kolmogorov-Smirnov Goodness-of-Fit Test

Sample size	**Level of Significance, α**				
n	**.20**	**.15**	**.10**	**.05**	**.01**
1	.900	.925	.950	.975	.995
2	.684	.726	.776	.842	.929
3	.565	.597	.642	.708	.828
4	.494	.525	.564	.624	.733
5	.446	.474	.510	.565	.669
6	.410	.436	.470	.521	.618
7	.381	.405	.438	.486	.577
8	.358	.381	.411	.457	.543
9	.339	.360	.388	.432	.514
10	.322	.342	.368	.410	.490
11	.307	.326	.352	.391	.468
12	.295	.313	.338	.375	.450
13	.284	.302	.325	.361	.433
14	.274	.292	.314	.349	.418
15	.266	.283	.304	.338	.404
16	.258	.274	.295	.328	.392
17	.250	.266	.286	.318	.381
18	.244	.259	.278	.309	.371
19	.237	.252	.272	.301	.363
20	.231	.246	.264	.294	.356
25	.210	.220	.240	.270	.320
30	.190	.200	.220	.240	.290
35	.180	.190	.210	.230	.270
over 35	$1.07/\sqrt{n}$	$1.14/\sqrt{n}$	$1.22/\sqrt{n}$	$1.36/\sqrt{n}$	$1.63/\sqrt{n}$