5 Cumulative Frequency Distributions, Area under the Curve & Probability Basics

5.1 Relative Frequencies, Cumulative Frequencies, and Ogives

We often have to compute the total of the observations in a class and all the classes before it (smaller in an ascending sort or larger in a descending sort). Figure 5-1 shows the *cumulative frequencies* for the ascending sort in column I.

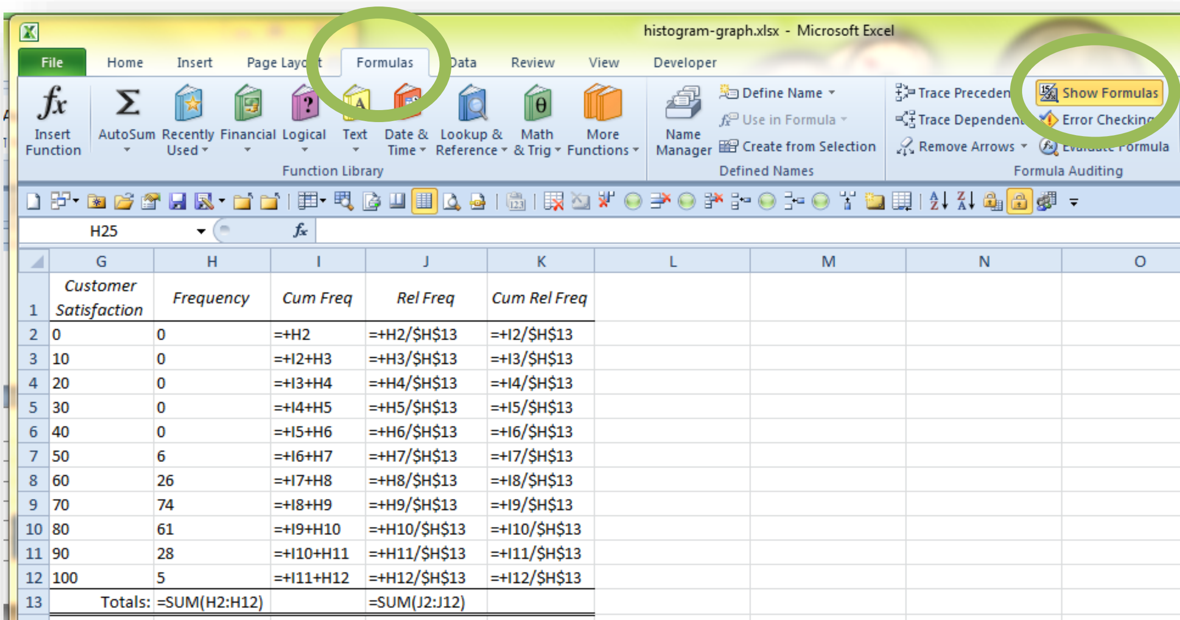
The proportion that a frequency represents in relation to the total of the frequencies (the sample size) is called a *relative frequency*. In Figure 5-1, the relative frequencies for the original distribution are shown in column J. The relative frequencies for the cumulative distribution are shown in column K.

The formulas for computing cumulative and relative frequencies are shown in Figure 5-2, which was generated by choosing the Formulas | Show Formulas buttons (green ovals in the figure).

Figure 5-1. Cumulative and relative frequencies.

G	H	I	J	K
Customer Satisfaction (0-100)	*Frequency*	*Cum Freq*	*Rel Freq*	*Cum Rel Freq*
0	0	0	0%	0%
10	0	0	0%	0%
20	0	0	0%	0%
30	0	0	0%	0%
40	0	0	0%	0%
50	6	6	3%	3%
60	26	32	13%	16%
70	74	106	37%	53%
80	61	167	31%	84%
90	28	195	14%	98%
100	5	200	3%	100%
Totals:	200		100%	

Figure 5-2. Formulas for cumulative and relative frequencies.

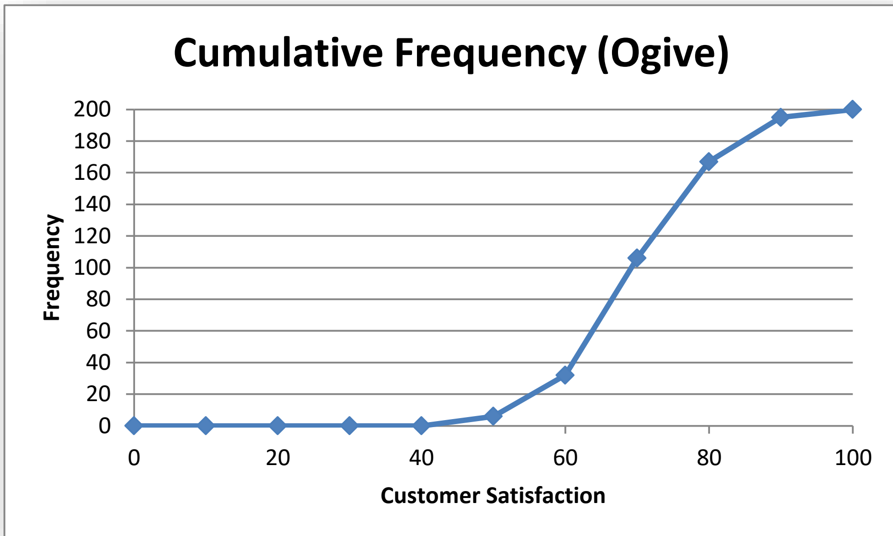


	G	H	I	J	K
1	*Customer Satisfaction*	*Frequency*	*Cum Freq*	*Rel Freq*	*Cum Rel Freq*
2	0	0	=+H2	=+H2/$H$13	=+I2/$H$13
3	10	0	=+I2+H3	=+H3/$H$13	=+I3/$H$13
4	20	0	=+I3+H4	=+H4/$H$13	=+I4/$H$13
5	30	0	=+I4+H5	=+H5/$H$13	=+I5/$H$13
6	40	0	=+I5+H6	=+H6/$H$13	=+I6/$H$13
7	50	6	=+I6+H7	=+H7/$H$13	=+I7/$H$13
8	60	26	=+I7+H8	=+H8/$H$13	=+I8/$H$13
9	70	74	=+I8+H9	=+H9/$H$13	=+I9/$H$13
10	80	61	=+I9+H10	=+H10/$H$13	=+I10/$H$13
11	90	28	=+I10+H11	=+H11/$H$13	=+I11/$H$13
12	100	5	=+I11+H12	=+H12/$H$13	=+I12/$H$13
13	Totals:	=SUM(H2:H12)		=SUM(J2:J12)	

The formula for the first cell in Column I for the *Cumulative Frequency* points at the first frequency (cell H2). The formula for the second cell in Column I (cell I3) is the sum of the previous cumulative frequency (I2) and the next cell in the *Frequency* column (H3). The $H means that propagating the formula elsewhere maintains a pointer to column H and the $13 freezes references so that propagating the formula maintains the pointer to row 13.

The line graph for a cumulative frequency is called an *ogive*. Figure 5-3 shows the ogive for the data in Figure 5-1.

Figure 5-3. Ogive for customer satisfaction data.



INSTANT TEST P 5-2

Find some frequency distributions with at least 20 categories in a research paper or statistical report in an area that interests you. Prepare two different frequency distributions based on different bins (e.g., 10 bins vs 20 bins) and create the charts that correspond to each. What are your impressions about using fewer and more categories (bins) in the representation of the frequency data?

5.2 Area under the Curve

One can plot the observed frequencies for the categories defined on the X-axis and examine the area under the curve.

Looking at Figure 5-4, the dark blue line represents the frequency of observations below any particular value of customer satisfaction. The area under the entire curve (shaded pale blue) represents the total number of observations – 200 in this example.

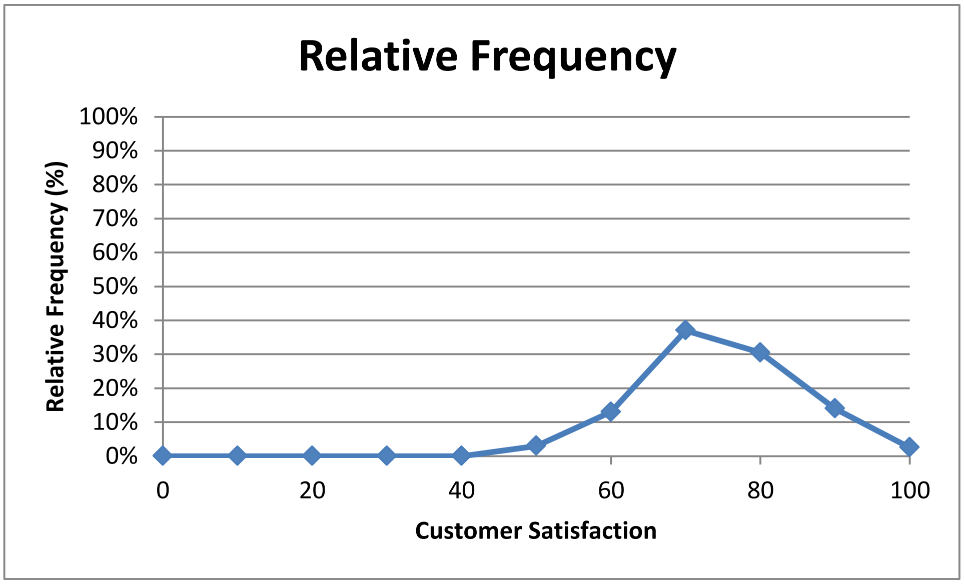
Figure 5-4. Frequency distribution for customer satisfaction scores.

In Figure 5-5, the pale green shaded area represents how many observations were below a specific value of Customer Satisfaction.

Figure 5-5. Areas under curve for observed frequencies.

If one constructs a graph of the frequency distribution with the *relative* frequency data, one ends up with a chart like Figure 5-6.

Figure 5-6. Area under the curve in a relative frequency distribution.



One of the most important principles in using frequency distributions is that the *area under the curve* represents the total proportion of all the observations in all the categories selected. Just as you can add up all the totals for each column in a group of adjacent categories in the histogram, the same principle applies to the frequency distribution.

To repeat, the area under the entire curve represents *100% of the observations*. The area to the *left* of a particular value on the X-axis (e.g., 80) represents the percentage of the observations from zero to just less than the value; e.g., for $X \leq 80$, the area represents the 84% of the total observations (see Column K in Figure 5-1).

5.3 Basic Concepts of Probability Calculations

By definition, an event that is absolutely certain has a probability of one (100%). For example, the probability that a person with a particular disease and specific health attributes will die within the next year is carefully estimated by statisticians working for insurance companies[55] to help set insurance rates. However, no one needs any sophisticated statistical knowledge to estimate that the likelihood that a person will eventually die some time in the future is 1: that's a known certainty.

Similarly, an impossible event has a probability of zero. Thus, unless one is living in a science fiction or fantasy story, the probability that anyone will be turned inside out during a transporter accident is zero.

If events are mutually incompatible (they can't occure at the same time) and they constitute all possibilities for a given situation, the sum of their individual probabilities is one. For example, a fair cubical die has a probability

$$p_i = 1/6$$

of landing with the top face showing one dot (**i**=1) facing up; indeed **p_i = 1/6** for all ***i***. The probability that a fair die will land showing a top face of either a 1 or a 2 or a 3 or a 4 or a 5 or a 6 is

$$\Sigma p_i = 1$$

This makes perfect sense, since the probability of something that is absolutely certain is by definition 1 – and if we exclude weird cases where a die balances on an edge, the only possible faces on the top of a six-sided die are 1, 2, 3, 4, 5, or 6.

If an event **i** has a probability **p_Y** then not having the event occur has

$$p_N = 1 - p_Y$$

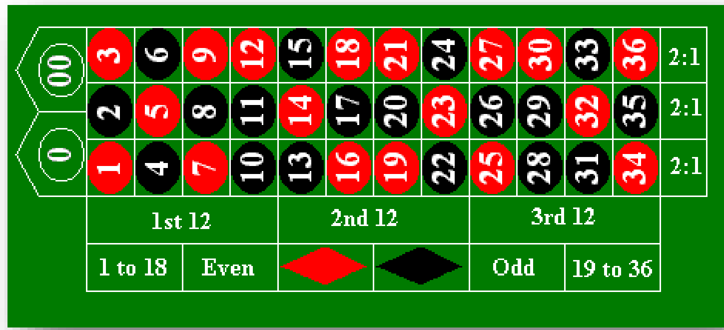
Figure 5-7. US casino roulette wheel.

- Thus the probability that a single throw of a fair die will *not* result with the 2-face upward is 1 – (1/6) = **5/6**. Another way of thinking about that is that there is one out of six ways of satisfying the description of the state and five out of six ways of not satisfying the description of the state.

[55] Statisticians who work on behalf of insurance companies are called *actuaries*.

- Similarly, since there are 38 slots on a standard roulette US wheel, as shown in Figure 5-7,[56] the probability that a gambler will win the 36:1 payment for placing a bet on a specific number (e.g., #18) is exactly **1/38**. The probability that the ball will *not* land on slot #18 is exactly 37/38.
- The probability that a roulette player will win the 2:1 payout for having a bet on the red box on the board when the ball lands in a red slot is exactly **18/38** and the probability that the ball will *not* land on a red slot is therefore 1 – (18/38) = **20/38**.

Figure 5-8. US casino roulette board.



If events are *independent* of each other (not influenced by each other), the the probability that two events p_1 and p_2 will occur at once (or in sequence) is

$$p_1 * p_2$$

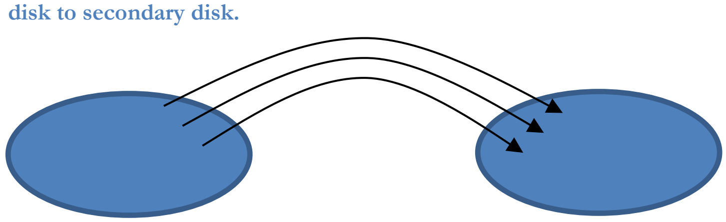
- For example, if you buy a lottery ticket with a 1 in 100,000 chance of winning $10,000, the chance of winning $10,000 is therefore 1/100,000. If you buy two of the same kind of lottery tickets, the chance of winning $10,000 on *both* of the tickets is (1/100,000) * (1/100,000) = 1/10,000,000,000 or simply 1e-5 * 1e-5 = **1e-10**.
- The probability of *losing* on *both* lottery tickets is (1 – 1e-5) * (1 – 1e-5) = 0.99999 * 0.99999 = 0.99998. The probability of winning on *at least one* ticket is 1 – 0.99998 = 0.00002.

This kind of reasoning is especially useful in calculating failure rates for complex systems. In information technology, a useful example of the probability-of-failure calculations is *Redundant Arrays of Independent Disks* – specifically RAID 1 and RAID 0 disk drives.

Here are the basics about these two configurations:

- RAID 1 (redundancy) improves resistance to disk failure (i.e., provides fault tolerance) by making bit-for-bit copies of data from a main drive to one or more mirror drives. If the main drive fails, the mirror drive(s) continue(s) to provide for I/O while the defective drive is replaced. Once the new, blank drive is in place, array management software can rebuild the image on the new drive. The frequency of mirroring updates can be defined through the management software to minimize performance degradation. As long as at least one of the disks is working, the array works.

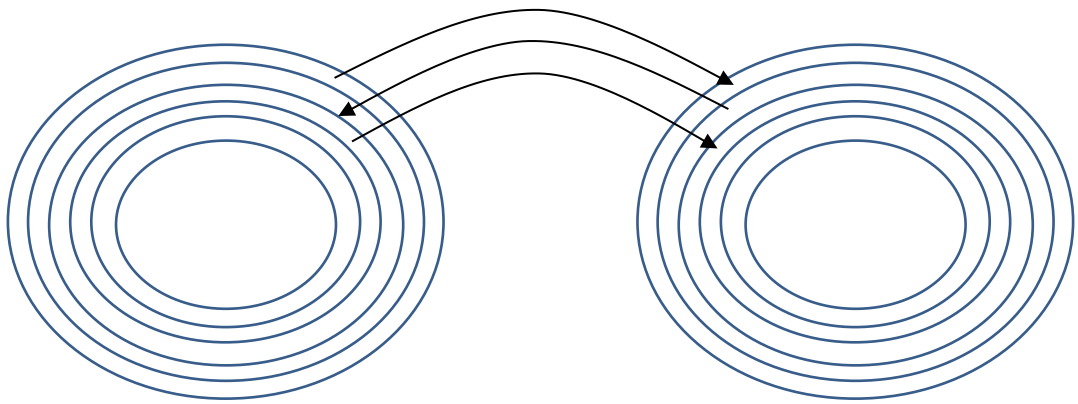
Figure 5-9. Raid 1 array with 2 disks showing writing from primary disk to secondary disk.



[56] There are 18 black slots and 18 red slots – all of which are involved in paying out money to the gambler depending on the bets – and two house slots (0 and 00) that result in having the house take all of the bets on the table without paying anything out.

- RAID 0 (speed) improves performance by *striping*, in which data are written alternately to cylinders of two or more disk drives. With multiple disk heads reading and writing data concurrently, input/output (I/O) performance improves noticeably. All the disk drives in the array must work for the array to work.

Figure 5-10. RAID 0 array showing writing to cylinders in alternate disks.



Now let's think through the probability of failure for each of these RAID configurations.

- Let the probability of failure of any one disk in a specified period be p (e.g., 1/100 per year = 0.01).
- For a RAID 1 (redundancy) array with n independent and identical disks, the probability that all n disks will fail is

$$P\{\text{all } n \text{ drives fail}\} = p^n$$

For example, with p = 0.01/year and two disks in the RAID 1 array, the chances that both disks will fail at the same time is only 0.01^2 or **0.0001** (one failure in a ten thousand arrays).

- For a RAID 0 (speed) array with n interleaved independent and identical disks, every disk must function at all times for the array to work. So first we compute the probability that all the drives will work.

$$P\{\text{all } n \text{ drives work}\} = (1 - p)^n$$

- For example, using the same figures as in the previous bullet, we compute that the chance of having both drives work throughout the year is $0.99^2 = 0.9801$.
- But then the chance that *at least one of the drives* will not work must be

$$P\{\text{at least one drive fails}\} = 1 - (1 - p)^n$$

and therefore the example gives a failure probability for the RAID 0 of 1 – 0.9801 = 0.0199 – almost double the probability of a single-disk failure.

- If there were 10 disks in the RAID 0, the probability of RAID failure using the same figures as the examples above would be

$$\mathbf{1 - (1 - 0.01)^{10} = 1 - 0.99^{10} = 1 - 0.9043821 = 0.0956}$$

or almost 10 times the single-disk failure.

The same kind of reasoning applies even if the probabilities of elements in a system are different. One then uses the following formulae:

- For redundant systems which work if *any* of the *n* components (p_1, p_2, $p_3 \ldots p_n$) in the calculation work, so that we need the probability that *all the components will fail,*

$$P\{\text{system failure}\} = p_1 * p_2 * p_3 * \ldots * p_n$$

which is more economically represented using the capital pi symbol (**Π**) for multiplication (much like the capital sigma (**Σ**) symbol for addition),

$$P\{\text{system failure}\} = \Pi p_i$$

- Similarly, for a redundant system,

$$P\{\text{system failure}\} =$$
$$1 - [(1 - p^1)*(1-p^2)*(1-p^3)*\ldots*(1-p^n)] =$$
$$1 - \Pi(1 - p_i)$$

INSTANT TEST P 5-8

Without referring to this text or to notes, explain the reasoning for how to calculate the probability that a two-disk RAID 1 (redundancy) array will fail within a year if the probability that a disk will fail in a year is known. Then explain how to calculate the probability of failure within a year for a RAID 0 (speed) array with three disks.

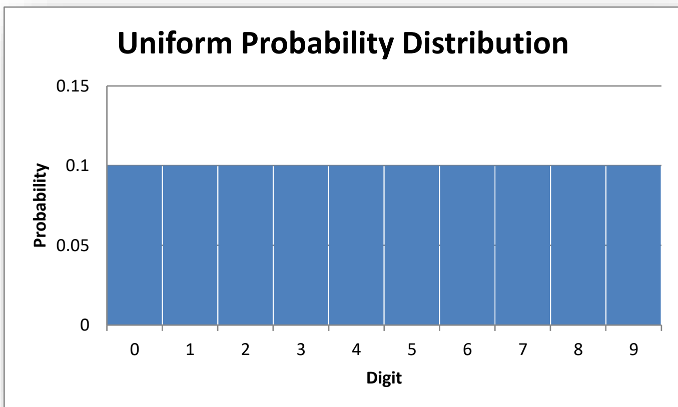
5.4 The Uniform Probability Distribution

As a simple example, suppose we count the number of each of the 10 digits (0, 1, 2…. 9) in an long series of numbers generated by a good random-number generator such as Excel's =RAND() or =RANDBETWEEN(lower, upper) functions. In the long run, we would expect 10% of all the digits to be 0, 10% to be 1, and so on. Figure 5-11 shows the *uniform probability distribution* corresponding to this thought-experiment. Each digit has a 10% (0.1) probability of being observed in a random sequence. We call this probality the *probability density function* and describe it as

$$P(x) = 0.1$$

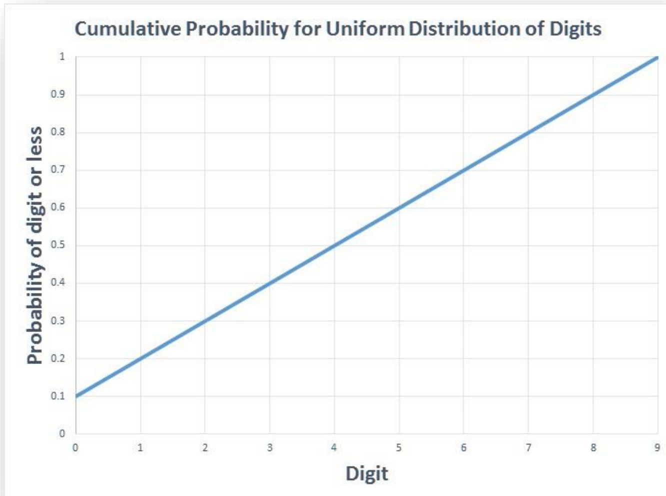
for all 10 values of x, the random deviate. Figure 5-11 shows the uniform distribution graphically.

Figure 5-11. Probability distribution for digits in random numerical data.



Now consider how many of the digits will be 3 or lower (3, 2, 1, 0): clearly there should be 40% (0.4) that correspond to that selection criterion. Accordingly, the *surface area under the curve* should be 40% – and that matches what we see in

Figure 5-12. Cumulative probability distribution for digits in random numerical data.



The area under the "curve" (in this case it's a straight line across the top of the rectangles) is simply

$$P(x \leq n) = 0.1(n + 1)$$

For example, the probability that a randomly selected digit will be 6 or less is 0.1(6+1) = 0.7, which matches what we can see on the graph. We call this function the *cumulative probability* function, often called *alpha* (α).

Finally, if we want to know how large the digit x_α is that includes α= 0.8 of the observations at or below its value, we can see on the graph that x = 7 satisfies the requirement. The function is notated as

$$x_\alpha = 10\alpha - 1$$

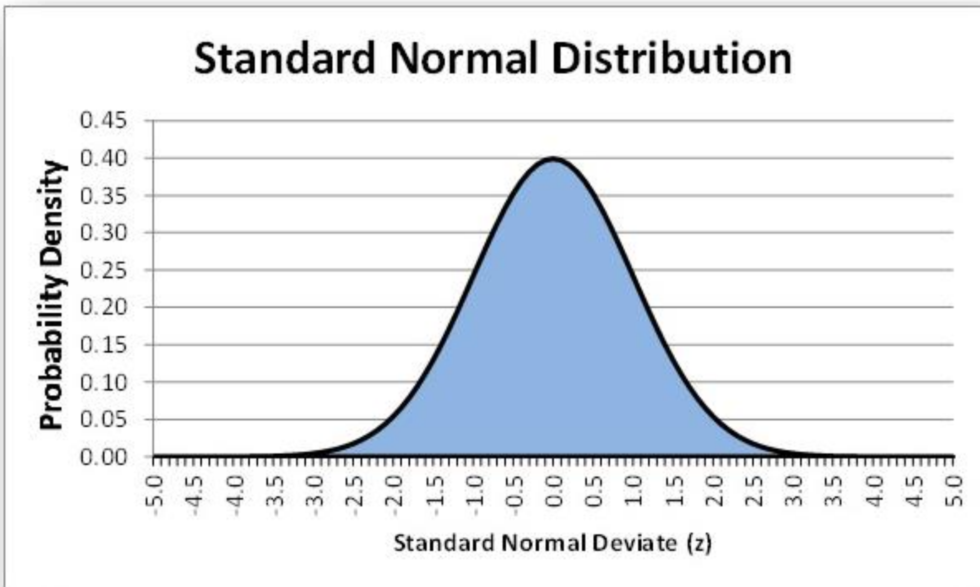
and is called the *inverse function* for the uniform distribution. Other distributions have their own formulas for calculating the probability density, the cumulative probability, and the inverse function. More important for our use, they are also computed for us automatically in EXCEL and in other statistical packages.

5.5 The Normal Probability Distribution

There are many theoretical probability distributions that we use in statistics. For example, the standard Normal distribution describes how observations vary when there are many factors that come together to generate the observation. A familiar example is peoples' weight: many factors contribute to the weight of an individual, and the *bell-curve*[57] shown in Figure 5-13 illustrates the kind of variation one would expect by computing the *standard Normal deviate*, *z*, based on the average (arithmetic mean) and on how much variability (measured by the standard deviation) there is in the population. We'll be studying these descriptive statistics in detail later.

In Figure 5-13, the dark bell-shaped curved *line* is the *probability density* function; this is a function that allows us to compute the probability that a deviate occurs in any given interval. Probability density functions are not probabilities: they often exceed the value 1.[58]

Figure 5-13. Standard Normal curve illustrating probability distribution.



The blue fill under the curve symbolizes the *area under the curve*; as always, the total is defined as 1 (100%) because all possible values of z are considered.

A probability of 1 means *certainty*; it is certain that the values of z must lie between -∞ and +∞; however, as you can see from the graph, the probability density function falls to infinitesimal levels around $z = -4$ and $z = +4$.

Half of this distribution lies at or below the mean and half falls above the mean.

---

[57] Also called a *Gaussian* distribution in honor of Carl Friedrich Gauss, who lived from 1777 to 1885 in Germany and who contributed greatly to probability theory, statistics, and many other aspects of science and mathematics. See (Gray 2012).

[58] For exhaustive details of the probability density function, see (Nykamp nd).

To find the exact values relating to the Normal distribution, we can use the Normal distribution functions in EXCEL, which several functions related to the Normal distribution, as shown in Figure 5-14. One can access this menu by typing =NORM in a cell.

These functions are explained in the HELP facility for EXCEL 2010, as shown in . The versions of the functions without the period separators are included in EXCEL for compatibility with older versions of EXCEL.

Figure 5-14. Normal distribution functions in Excel 2010.

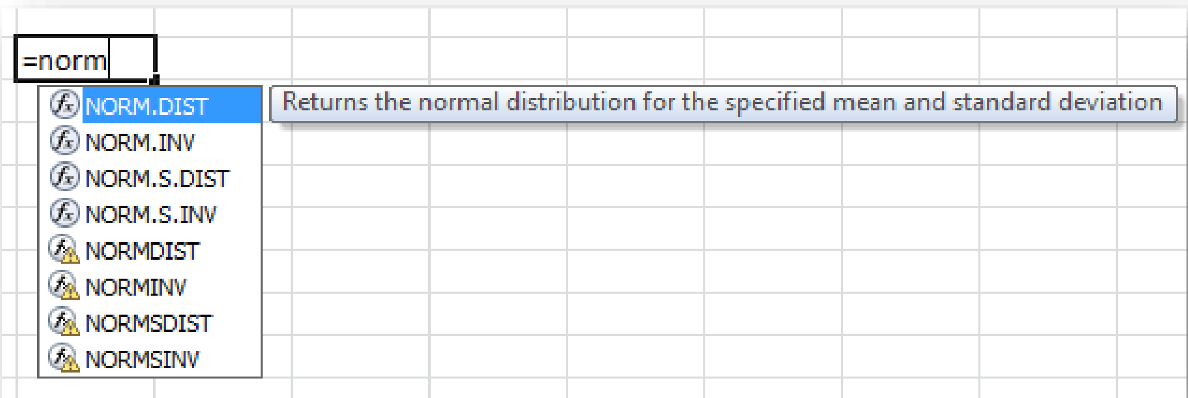
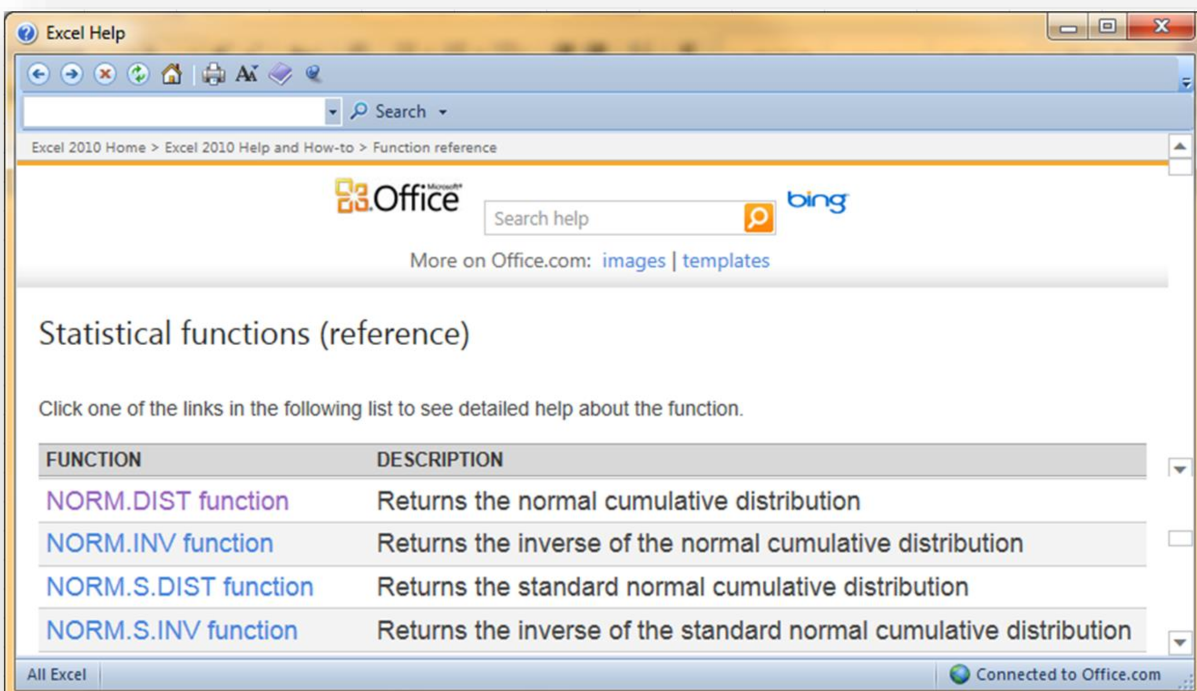


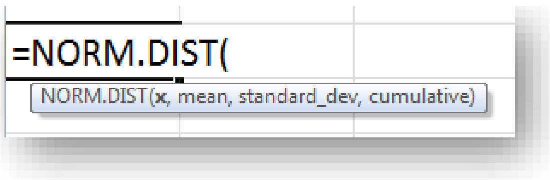
Figure 5-15. Excel 2010 HELP for Normal distribution functions.



5.6 Area under the Curve for Any Normal Distribution

In applied statistics, you will constantly have to find the area under a theoretical probability distribution curve. The area of interest is generally symbolized by lower-case alpha (α). For the Normal distribution, the =NORM.DIST (=NORMDIST in the older versions of EXCEL) function returns (calculates) the area of the curve to the *left* of a specified Normal deviate. For example, typing =NORM.DIST (or left-clicking on that function in the menu shown in Figure 5-14 instantly brings up the *parameters* required for the function to work, as shown in Figure 5-16.

Figure 5-16. Pop-up reference to parameters for =NORM.DIST function in Excel 2010.



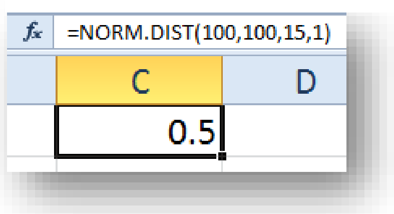
The first parameter, x, is the value of interest; mean is the average of the distribution, standard_dev is the standard deviation (σ, pronounced *sigma*, which we will study in detail later) and cumulative is 0 for the *probability density function* and 1 for the *cumulative probability* (area of the curve to the left of) for x.

Intuitively, if we have a Normal distribution such as the IQ distribution, which theoretically has mean IQ = 100 and standard deviation = 15, half the values will lie at or below the mean. So calculating

=NORM.DIST(100,100,15,1)

should give us $a = 0.5$. And sure enough, it does, as shown in Figure 5-17.

Figure 5-17. Calculating a cumulative probability value using NORM.DIST in Excel 2010.



Given the appearance of the bell curve and the logic of probabilities for mutually exclusive cases, the probability that a random x will be greater than the mean for this example is given by

$$P(x > 100) = 1 - P(x \leq 0) = 1 - 0.5 = 0.5$$

INSTANT TEST P 13

Practice using the =NORM.DIST function with different values to ensure that you become familiar with it. As a check on your work you can remember that a Normal distribution with mean = 0 and standard_dev = 1 has 0.025 to the left of x = -1.96 and 0.975 to the left of x = +1.96. Play around with different values of the variables. You may want to *point* to cells for the variables you are changing rather than having to modify constants in the formula itself. Be sure to experiment with the cumulative parameter to see for yourself what it does.

5.7 Area Under the Curve for the Standard Normal Distribution

The standard Normal deviate *z* is

$$z = (Y - \mu)/\sigma$$

where

- Y is a value for the abscissa
- μ is the mean of the population
- σ is the parametric standard deviation of the population.

The =NORM.S.DIST function generates the probability that random observation from a Normal distribution will generate a standardized Normal deviate at or below the indicated value *z* as shown in Figure 5-19.[59]

The second parameter, cumulative, is 1 for the cumulative distribution and 0 for the (rarely-used) probability density function.

Figure 5-19. =NORM.S.DIST function in Excel 2010.

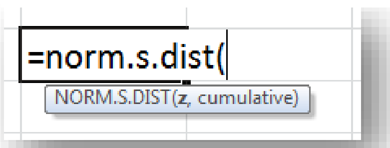
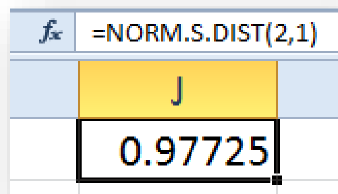


Figure 5-18. Calculation of probability that z ≤ 2 in Excel 2010.



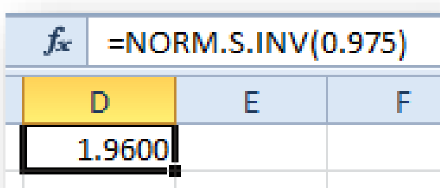
For example, the area under the curve at and to the left of z = 2 is 0.97725, as shown in Figure 5-18. In other words, the probability of observing a z ≤ 2 by chance in a standardized Normal distribution is 97.725%.

The =NORM.S.INV function provides the *critical value* of z corresponding to the α for the left tail of the distribution. How big must z be to demarcate a given left-tail α?

Figure 5-20 shows that the value of z = 1.96 demarcates a left-tail area under the standard Normal curve of 0.975. Thus the area on the right must be 1 – 0.975 = 0.025.

Because the Normal curves are symmetrical, this finding also implies that 0.025 of the curve lies to the *left* of -1.96. And indeed the function =NORM.S.INV(0.025) does generate -1.9600 (not shown).

Figure 5-20. Finding the critical value for a left-tail area in the standard Normal curve using an Excel 2010 function.



[59] By accident, the function name was typed in lowercase before this screenshot was taken. Uppercase and lowercase don't matter when typing function names in Excel.

5.8 Using EXCEL Functions for Areas Under Other Probability Distribution Curves

In addition to the Normal distribution, other statistical probability distributions that you will use extensively are the

- Chi-square (χ^2): important for *goodness-of-fit tests* (seeing if sampled data fits a theoretical expected distribution) and *tests of independence* (seeing if the underlying proportions of different classes are the same in different samples);
- F: for comparing the underlying *variances* (measure of variability) of two samples; particularly important in the valuable *analysis-of-variance* (*ANOVA*) methods;
- Student's t: for comparing the underlying statistics from different samples to see if they are the same; also for computing *confidence limits* (estimates of what the statistics could actually be given the sample values) to a wide variety of statistics.

The basic steps for learning about a statistical function in any computational package are to

- Find the functions by using a reasonable identifier – in EXCEL, one can start by typing out =name (e.g., =norm) where name represents the name of the distribution;
- Learn what the parameters are (in EXCEL, check the pop-ups and use HELP);
- Try computing something for which you know the answer to check that you're doing it right (e.g., use statistical tables for known values and see if you can duplicate the results);
- You can also compute complementary values to check that you're using the right function – e.g., calculating =norm.s.dist(0.05) = -1.644854 and =norm.s.dist(0.95) = 1.644854.

Some commonly used terminology for statistical functions includes

- *Left-tailed probability*: likelihood of observing a value *smaller than or equal to* the test value if the assumptions of the statistical function are true;
- *Right-tailed probability*: likelihood of observing a value *larger* (or sometimes *larger than or equal to*) than the test value if the assumptions of the statistical function are true.

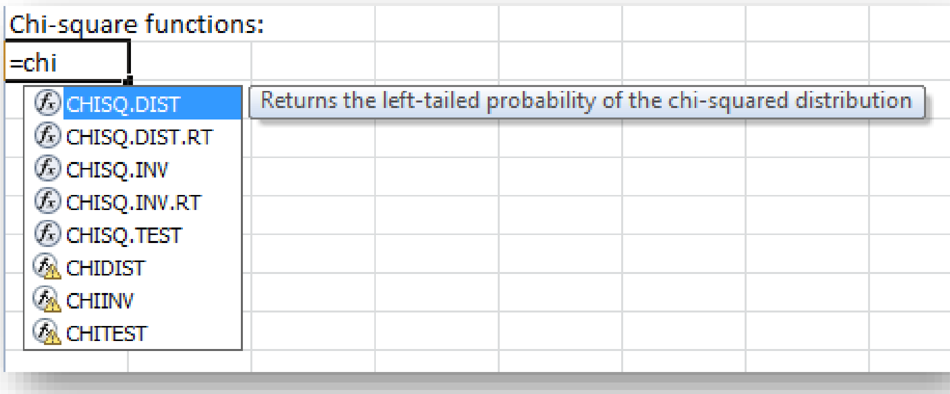
In general, there is little or no difference in practice between the area of a probability distribution to the *left of a value* ($< x$) and the area to the *left of or equal to a value* ($\leq x$).

5.9 Chi-Square Distribution

Typing =chi into a cell in EXCEL 2010 brings up a menu of choices, as shown in Figure 5-21.

Clicking on any one of the functions generates a list of its parameters, as shown in Figure 5-22. The middle

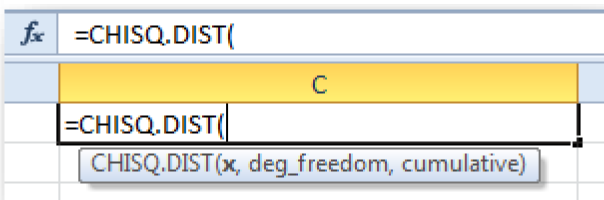
Figure 5-21. Menu of chi-square functions in Excel 2010.



term, deg_freedom, refers to *degrees of freedom* which are a function of the particular application of the chi-square.

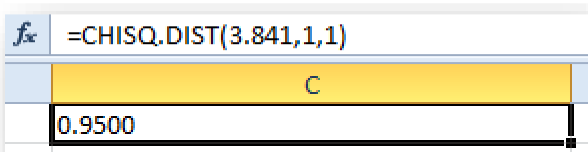
To check that you are using the =CHISQ.DIST function correctly, you can remember that 95% of the curve

Figure 5-22. Left-tail chi-square probability function in Excel 2010.



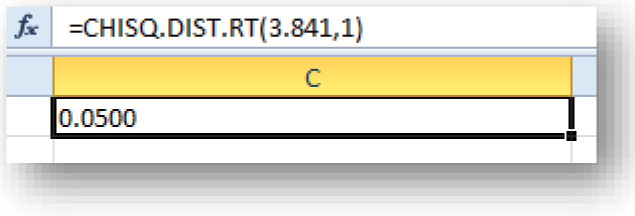
lies to the left of 3.841 when the distribution has 1 *degree of freedom*, as shown in Figure 5-23.

Figure 5-23. Left-tail chi-square probability function in Excel 2010.



That value also implies that 5% of the curve (1-0.95) lies to the *right* of 3.841 for 1 degree of freedom, as confirmed in Figure 5-24 using the EXCEL 2010 function =CHISQ.DIST.RT.

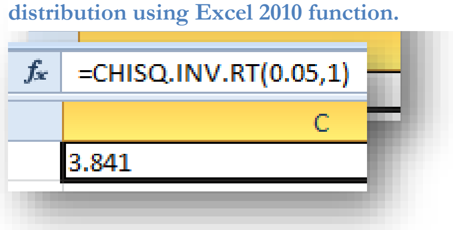
Figure 5-24. Right-tail chi-square probability function in Excel 2010.



The inverse functions =CHISQ.INV and =CHISQ.INV.RT produce the critical values for left-tailed and right-tailed probabilities, respectively. Thus for a left-tailed α = 0.95 with one degree of freedom, the critical value is 3.841, as shown in Figure 5-25.

Figure 5-25. Critical value for left-tail = 0.95 in chi-square distribution using Excel 2010 function.

Figure 5-26. Critical value for right tail = 0.05 in chi-square distribution using Excel 2010 function.



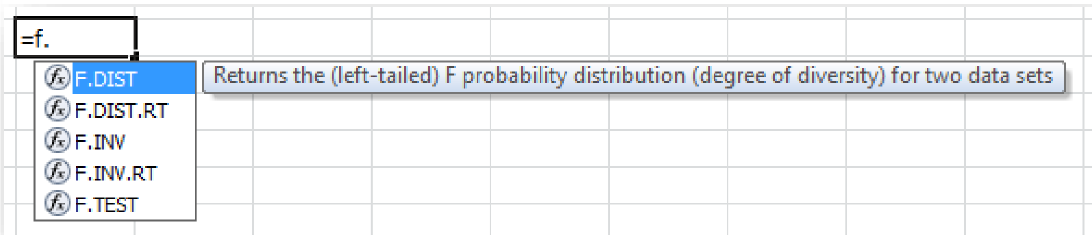
That finding also means that the right tail must be 0.05, and that's confirmed also, as shown in Figure 5-26.

We'll be looking at the =CHISQ.TEST function later in the course, where it plays an important role in testing *frequency distributions* for *categorical* data against theoretical distributions – tests of *goodness of fit* and of *independence*.. It actually calculates a value of a sample chi-square that can be tested against the theoretical distribution to see how likely it is that such a large value could occur by chance alone if the assumptions of the test were correct. The =CHITEST function, preserved in EXCEL for compatibility with older versions of EXCEL, generates the same result.

5.10 F Distribution

The F-distribution functions in EXCEL 2010 are shown in Figure 5-27. F-tests are critically important in analysis of varance (ANOVA) methods widely used in applied statistics.

Figure 5-27. F-distribution functions in Excel 2010.



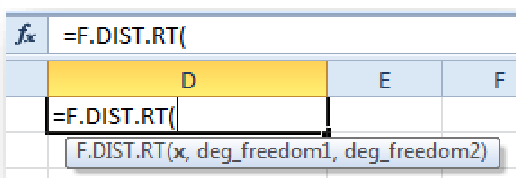
The HELP facility documents the functions, as shown in Figure 5-28.

Figure 5-28. HELP for F-distribution functions in Excel 2010.

F.DIST function	Returns the F probability distribution
F.DIST.RT function	Returns the F probability distribution
F.INV function	Returns the inverse of the F probability distribution
F.INV.RT function	Returns the inverse of the F probability distribution
F.TEST function	Returns the result of an F-test

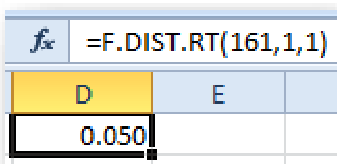
Each function is documented in the HELP facility by clicking on the link in Figure 5-28. In addition, starting to type the function in an EXCEL cell brings up a pop-up reminding the user of the parameters required, as shown in Figure 5-29.

Figure 5-29. Pop-up menu in Excel 2010 for right-tail area under the curve for F distribution.



For reference, you can note that the value F = 161 cuts off 5% of the distribution curve for 1 and 1 degrees of freedom, as shown in Figure 5-30.

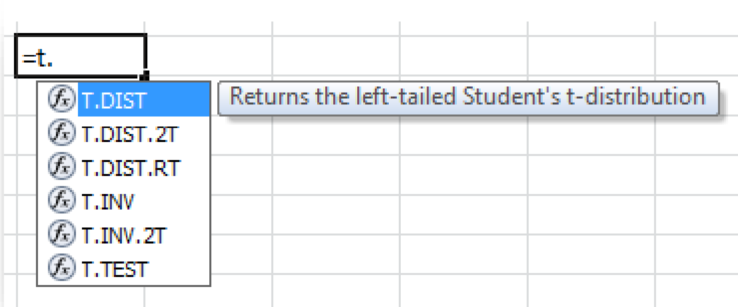
Figure 5-30. Critical value for right-tailed probability of 0.05 in F distribution with 1 & 1 df in Excel 2010.



5.11 Student's-t Distribution

In EXCEL 2010, typing =T. (or =t.) brings up the menu of *Student's t* [60] functions as shown in Figure 5-31. The =T.DIST function generates the left tail of the distribution. For example, t = 12.706 for a distribution

Figure 5-31. Student's t functions in Excel 2010.



with 1 degree of freedom cuts off 97.5% to the left of that value, as shown in Figure 5-32. The symbol .2T in the name of the function indicates a *two-tailed probability*, as shown in Figure 5-33. The symbol .RT in the function name indicates a right-tailed probability, as shown in Figure 5-34.

Figure 5-32. Left-tail probability for t distribution in Excel 2010.

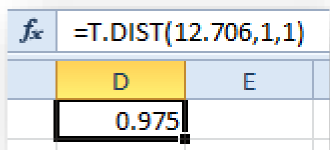


Figure 5-33. Two-tailed probability for t distribution in Excel 2010.

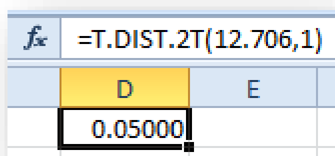
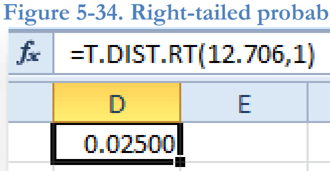


Figure 5-34. Right-tailed probability for t distribution in Excel 2010.



[60] Described in English in 1908 by William S. Gosset, who published under the pseudonym *Student*. See (Encyclopaedia Britannica 2012).

