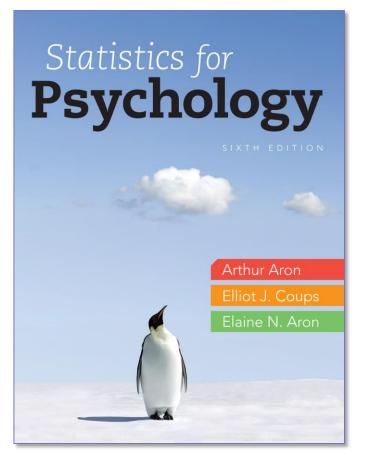
Statistics for **Psychology**

SIXTH EDITION





Some Key Ingredients for Inferential Statistics

 Statistics for Psychology, Sixth Edition

 ALWAYS LEARNING

 Arthur Aron | Elliot J. Coups | Elaine N. Aron

 Copyright © 2013 by Pearson Education, Inc. All Rights Reserved

PEARSON

Some Key Ingredients for Inferential Statistics

- Psychologists conduct research to test a theoretical principle or the effectiveness of a particular procedure
- The measurements are used to test a theory
- Researchers use inferential statistics to make more general conclusions about the theoretical principle
- These conclusions go beyond the particular group of research participans studied



Describing a Score

- After the measurement is done, researchers have some data: a group of scores from the research participants
- Knowing one score tells little about how it relates to the whole distribution of scores
- Comparing a score to the mean of a distribution does indicate whether a score is above or below average
- So, researchers need to describe a particular score in terms of where it fits into the overall group of scores



Relating a Score to the Mean

- Knowing the standard deviation of a distribution indicates
 - how much above or below average a score is in relation to the spread of scores in the distribution



- I ask Ayşe « to what extent are you a morning person?
- 1-7 point-scale (1 not at all; 7 extremely)
- She answers 5
- If I do not know anything a/ how other sts answer this question, I cannot say whether Ayşe is more or less of a morning person in relation to other sts
- If I know that M is 3.40 and SD is 1.47 in general for this question, Ayşe is above average
 - **5-3.40**= 1.60



Z Scores

- Z score makes use of the mean and SD to describe a particular score
- Number of standard deviations a score is above or below the mean
- Formula to change a raw score to a Z score: $Z = \frac{X - M}{SD}$
- If actual score is above the mean; Z is positive
- If actual score is below the mean; Z is negative



Z Scores

$$Z = \frac{X - M}{SD}$$

- For Ayşe example,
- Z=(5-3.40)/1.47= 1.09
- Ayşe's score is a little more than 1 SD above the mean
- Suppose Ömer answered 2
- Ömer's Z score is
 - Z=(2-3.40)/1.47= -.95
 - Ömer's score is a little less than 1 SD below the mean
 - Ömer's score is below the average by about as much as sts typically vary from average



Z Scores -2

Formula to change a Z score to a raw score:

$$X = (Z)(SD) + M$$



Steps to change a raw score to a Z score

- 1. Figure the deviation score
 - 1. Subtract the mean from the raw score
- 2. Figure the Z score
 - 1. Divide the deviation score by the standard deviation



- In a sample with a sample mean of 25 and a standard deviation of 5, calculate the z-score for the raw score 32.
 Z=(X-M)/SD
 = (32-25)/5
 - =+1.40

- In a sample with a sample mean of 50 and a standard deviation of 10, calculate the z-score for the raw score 44.
 - Z = (X-M)/SD
 - = (44-50)/10
 - = -0.60

• In a population with a mean of 100 and a standard deviation of 16, calculate the z-score for the raw score 132.

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

$$Z = (132 - 100) / 16 = +2.00$$

Computing a Raw Score

• When a *z*-score is known, this information can be used to calculate the original raw score. The formula for this is

$$X = (Z)(SD) + M$$
$$X = (z)(\sigma x) + \mu$$

• In a sample with a sample mean of 25 and a standard deviation of 5, calculate the raw score for z = -0.43

$$X = (Z)(SD) + M$$

$$= (-0.43)(5) + 25$$

= 22.85

• In a sample with a sample mean of 50 and a standard deviation of 10, calculate the raw score for z = -1.30

$$X = (Z)(SD) + M$$

$$= (-1.30)(10) + 50$$

=37

• In a population with a mean of 100 and a standard deviation of 16, calculate the raw score for z = +1.40 $X = (z)(\sigma x) + \mu$ X = (Z)(SD) + M= (1.40)(16)+100= 122.40

The Mean and SD of Z scores

- The mean of the distribution of Z scores is always zero.
 - This is why when you change a raw score to a Z score, you take the raw score minus the mean
 - So the mean is subtracted out of all the raw scores, making the overall mean come out to zero.
 - In other words; in any distribution, the sum of positive Z scores must always equal the sum of the negative Z scores

The Mean and SD of Z scores

- The SD of a set of any distribution of Z scores is always 1
 - This is when you change each raw score to a Z score, you divide the score by one SD



Z Score

- Z score is sometimes called as a standard score
- Two reasons for this,
 - **1.** Z scores have standard values for the mean and the SD
 - 2. Z scores provide a kind of standard scale of measurement for any variable

PEARSON

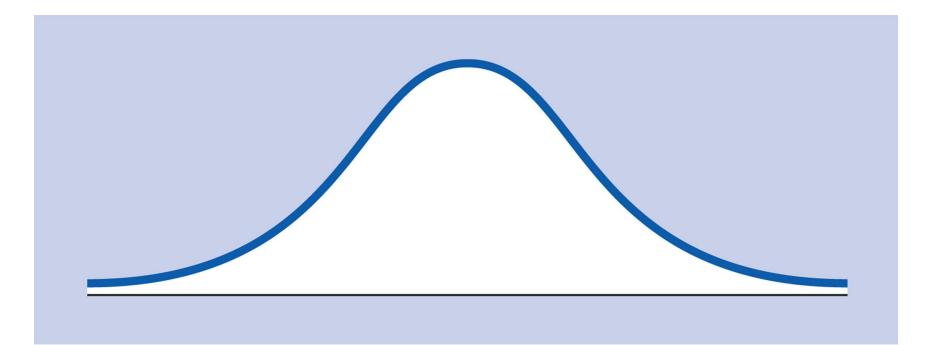
The Normal Curve

- In chapter 1, we covered that the garphs of the distributions of many of the variables that psychologists study follow a unimodal, roughly symmetrical, bell-shaped curve
- These bell shaped smooth histograms approximate a precise and important mathematical distribution called the normal distribution
- More simply, the normal curve



The Normal Curve

Normal curve



 Statistics for Psychology, Sixth Edition

 ALWAYS LEARNING

 Arthur Aron | Elliot J. Coups | Elaine N. Aron

 Copyright © 2013 by Pearson Education, Inc. All Rights Reserved



Why the normal curve is so common in nature

- Example: the number of letters a particular person can remember
- On some testings the number of letters remembered may be high, on others low and on most smw in b/w
- The number of different letters that a person can recall on various testings probably approximately follows a normal curve
- If a person has an ability to recall seven letters in this kind of memory task, on any particular testing, the actual number recalled will be affected by various influences
 - Noisiness of the room
 - The person's mood etc.



Why the normal curve is so common in nature

- The particular combinations of such influences that come up at any testing is essentially random,
- Thus, on most testings, positive and negative influences should cancel out
- This creates a unimodal distribution that is symmetrical
 - Because the number of letters recalled is as likely to be above as below the middle
- Being a unimodal, symmetrical curve does not guarantee that it will be a normal curve
 - It could be too flat or too pointed



Why the normal curve is so common in nature

- However, it can be shown mathematically that in the long run,
 - if influences are truely random AND
 - The number of different influences being combined is large;

A precise normal curve will result

Mathematical staticians call this principle the *«central limit theorem»*



Normal Distribution Characteristics

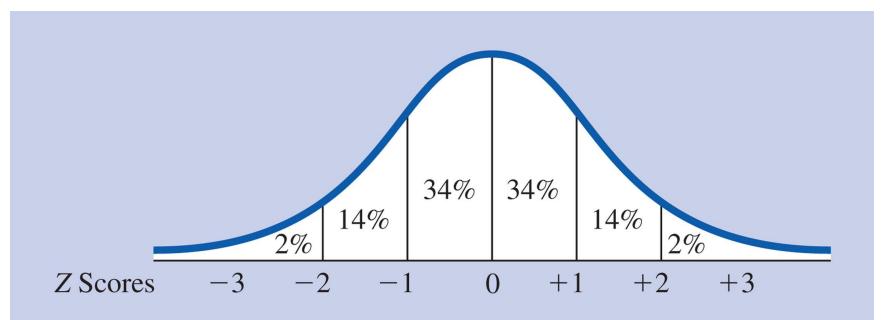
- Symmetrical
- Unimodal
- Most scores fall near center, fewer at extremes



- The shape of the normal curve is standard
- There is known percentage of scores above or below any particular point
- Exactly 50% of the scores in a normal curve are below the mean, because in any symmetrical distribution half the scores are below the mean
- Let's see what percentages fall between mean and 1 SD and between 1 SD and 2 SD;



 Normal curve and percentage of scores between the mean and 1; and 2 standard deviations from the mean



 Statistics for Psychology, Sixth Edition

 ALWAYS LEARNING
 Arthur Aron | Elliot J. Coups | Elaine N. Aron

 Copyright © 2013 by Pearson Education, Inc. All Rights Reserved

PEARSON

- E.g., IQ
- On many widely used intelligence tests, the mean IQ is 100; the SD is 15 and the distribution of IQs is roughly normal curve
- By knowing these; we can infer that 34 % of people have IQs between 100 and 115
 - 100 (the mean) and 115 (Iq score that is 1 SD above the mean)
- Percentage b/w 1 SD and 2 SD is 14% which is lower than the percentage b/w the mean and 1 SD
 - Fewer scores in between 1 SD and 2 SD than those in b/w the mean and 1 SD



- E.g., IQ
- Percentage from 2 SD and above is approximately 2 %
 - Very fewer scores from 2 SD and above



The normal curve table and Z scores

- When we want to use more precise percentages?
 - 10 % in tail \rightarrow corresponds to what Z score?
 - b/w mean and which Z score \rightarrow 23 %
- The normal curve table shows the precise percentage of scores between the mean (Z score of 0) and any other Z score
- Table also includes the precise percentage of scores in the tail of the distribution for any Z score
- Table lists positive *Z* scores
 - Since distribution is symmetrical, percentages in tails are also symmetrical



		A	A	
	z	% Mean to Z	% in Tail	
			50.00	
			49.60	
			49.20	
		1.20	48.80	
	.04		48.40	
			48.01	
		2.39	47.61	
		2.79	47.21	
		3.19	46.81	
	.09		46.41	
	.10		46.02	
	.11	4.38		
	.12	4.78		
ti: nu Y	.13	5.17	44.92	

PEARSON

Figuring Areas under the Normal Distribution

- Steps for figuring percentage area above or below particular raw or Z score:
 - 1. Convert raw score to Z score (if necessary)
 - 2. Draw normal curve, locate where Z score falls on it, and shade in area for which finding percentage
 - 3. Make rough estimate of shaded area's percentage (using 50%-34%-14% rule)



Figuring Areas under the Normal Distribution

- Steps for figuring percentage area above or below particular raw or Z score:
 - 4. Find exact percentage using normal curve table
 - 5. If needed, add or subtract 50% from this percentage
 - 6. Check that exact percentage is within range of estimate from Step 3



Figuring Z Scores and Raw Scores from Percentages

- Steps for figuring Z scores and raw scores from percentages:
 - Draw normal curve, shade in approximate area for the percentage (using the 50%-34%-14% rule)
 - 2. Make rough estimate of *Z* score where shaded area stops
 - 3. Find exact Z score using normal curve table



Figuring Z Scores and Raw Scores from Percentages -2

- Steps for figuring Z scores and raw scores from percentages:
 - Check that Z score similar to rough estimate from Step 2
 - 5. To find a raw score, change it from Z score



z	% Mean to Z	% in Tail	z	% Mean to Z	% in Tall
.00	a 00	50.00	.30	11.79	38.21
.01	0.40	49.60	.31	12.17	37.83
.02	0.80	49.20	.32	12.55	37.45
.03	1.20	48.80	.33	12.93	37.07
.04	1.60	48.40	.34	13.31	38.69
.05	1.99	48.01	.35	13.68	36.32
06	2.39	47.61	.36	14.06	35.94
.07	2.79	47.21	.37	14.43	35.57
63	3.19	46.81	.38	14.80	35.20
.09	3.59	46.41	.39	15.17	34.83
10	3.98	46.02	.40	15.54	34.46
.11	4.38	45.62	.41	15.91	34.09
12	4.78	45.22	.42	16.28	33.72
13	5.17	44.83	.43	16.64	33.36
14	5.57	44.43	.44	17.00	33.00
15	5.96	44.04	.45	17.36	32.64
16	6.36	43.64	.46	17.72	32.28
17	6.75	43.25	.47	18.08	31.92
.18	7.14	42.86	.48	18.44	31.56
.19	7.53	42.47	.49	18.79	31.21
20	7.93	42.07	.50	19.15	30.85
21	8.32	41.68	.51	19.50	30.50
22	8.71	41.29	.52	19.85	30.15
.23	9.10	40.90	.53	20.19	29.81
24	9.48	40.52	.54	20.54	29.46
.25	9.87	40.13	.65	20.88	29.12

Table 0-1 Normal Curve Areas: Percentage of the Normal Curve Between the Mean and the

Statistics for Psychology, Sixth EditionALWAYS LEARNINGArthur Aron | Elliot J. Coups | Elaine N. AronCopyright © 2013 by Pearson Education, Inc. All Rights Reserved

PEARSON

- IQ scores M=100 and SD=15
- If a person has an IQ of 125, what percentage of people have higher IQs?
- 1. Z = (125 100)/15 = 1.67
- Draw a picture of the normal curve and decide where the Z score falls on it and shade the area for which you are finding the percentage
- 3. Make a rough estimate of the shaded area
- 4. Find the exact percentage by using the normal curve table
- Check this result w/ step 3 estimate (your rough estimate)



- IQ scores M=100 and SD=15
- What IQ score would a person need to be in the top 5%?
- 1. Draw normal curve, shade in approximate area for the percentage (using the 50%-34%-14% rule)
- 2. Make rough estimate of Z score where shaded area stops
 - 1. b/w 1 SD and 2 SD
- 3. Find exact Z score using normal curve table
 - 1. It is 1.64
- 4. Check that Z score similar to rough estimate from Step 2
 - 1. 1.64 is b/w 1 and 2 ☺
- 5. To find a raw score, change it from Z score

1. X = Z.SD + M = 1.64x15 + 100 = 124.60

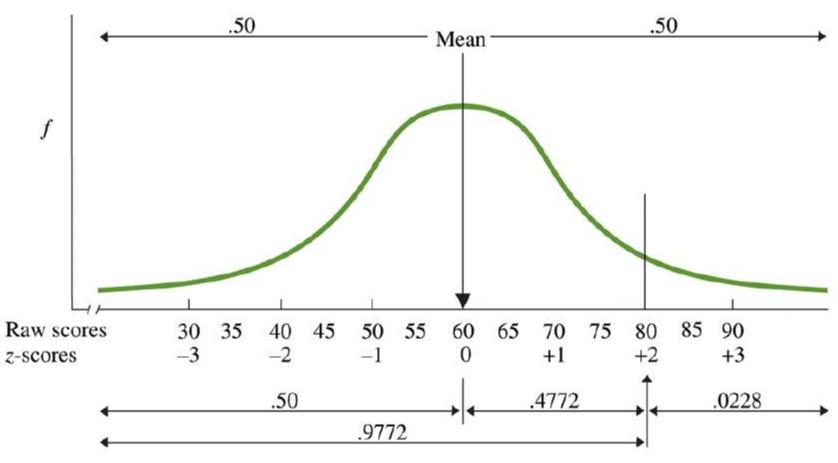
To be in the top 5 %, a person would need an IQ of at least 124.60

Statistics for Psychology, Sixth Edition ALWAYS LEARNING Arthur Aron | Elliot J. Coups | Elaine N. Aron Copyright © 2013 by Pearson Education, Inc. All Rights Reserved

- What range of IQ scores includes the 95% of people in the middle range of IQ scores?
- Z = +/- 1.96
- Apply formula:
 - X = 70.60 and X= 129.40
- The middle 95% of IQ scores run from 70.60 and 129.40



Proportions of the Standard Normal Curve



- In a sample with a sample mean of 40 and a standard deviation of 4, find the relative frequency of the scores above 45.
 - Z=(X-M)/SD= (45-40)/4 = +1.25

Above 1.25 is equal to =.1056 from the table.



- In a sample with a sample mean of 40 and a standard deviation of 4, find the percentile of the score 41.5
 - Z=(X-M)/SD
 - = (41.5-40)/4
 - = +0.38

Area between the mean and 41.5 is .1480 from the table.

.50+.1480=.6480 -> 64.801 is 65th percentile 24

- In a sample of 1000 scores, with a mean of 65 and a standard deviation of 12,
 - a) What is the percent of scores below 59?
 - b) How many scores are between the mean and 70?
 - c) Which raw score signifies the top 3%?



a) What is the percent of scores below 59?

z = (59-65)/12 = -0.50

Below means the lower tail, so we look at column C and find 30.85







b) How many scores are between the mean and 70?

• Simple frequency:

z = (70-65)/12 = +0.42

To find the percent between a score and the mean we look at column B and find 16.28

How many of the scores are between the mean (65) and 70?

We have 1000 scores; so, (1000/100) * 16.28 = 162.8 = almost163 of the scores are between 65 and 70





Example 6

- c) Which raw score signifies the top 3%?
- Raw score:
 - Because it says top 3% we look at column C and find the raw score with the proportion that is closest to 3.00
 - It is the score 3.01 with a z-score of +1.88
 - $-X = (z) * (S_D) + M$ -X = 1.88 * 12 + 65-X = 87.56





- What IQ score would a person need to be in the top 55%?
 - 1. Draw a picture of the normal curve and shade in the approximate area for your percentage by using the 50%-34%-14& percentages. There are 50% of scores above the mean. There are 34% scores btw. the mean and 1 SD below the mean, so the score is btw. the mean and 1 SD below the mean.

PEARSON

- 2. Make rough estimate of the Z score where the shaded area stops. Z score has to be btw. 0 and -1.
- Solution
 Find the exact Z score using the normal curve table. Being in the top 55% means that 5% of people have IQs btw. this IQ and the mean (55%-50%=5%)
- In the normal curve table, closest percentage to 5% is 5.17%, which goes with Z score of .13. Because you are below the mean, this becomes -.13.



Example 8a

- Suppose a test of sensitivity to violence is known to have a mean of 20, a SD of 3, and a normal curve shape.
- a) What percentage of people have scores above 24?
- **1.** Change a raw score to a Z score.
- Z= (X-M)/SD Z=(24-20)/3= 1,33
- 2. Draw a picture of normal curve.



Example 8a

 3. Make a rough estimate of the shaded area's percentage based on the 50%-34%-14& percentages. If the shaded area started at a Z score of 1, it would include 16%. If the shaded area started at a Z score of 2, it would include only 2%. So, it has to be somewhere btw. 16% and 2%.



Example 8a

 4. Find the exact Z score using the normal curve table. In the table A-1, 1.33 in the Z column goes with 9.18%.



- Suppose a test of sensitivity to violence is known to have a mean of 20, a SD of 3, and a normal curve shape.
- a) What minimum score of a person needs to be in the top 75%?



- 1. Draw a picture of the normal curve and shade in the approximate area for your percentage by using the 50%-34%-14& percentages. The shading has to begin btw. the mean and 1 SD below the mean.
- 2. Make rough estimate of the Z score where the shaded area stops. Z score has to be btw. 0 and -1.



- **3.** Find the exact Z score using the normal curve table.
- 75%-50%=25% (so, you need to include 25% below the mean)
- Looking in the '% Mean to Z' column of the normal curve table, closest figure to 25% is 24.86, which goes with a Z of .67. Since we are interested in below the mean, we want -.67.



- 4. Check your exact Z score is similar to your rough estimate from Step2. -67 is btw. 0 and -1.
- **5.** If you find a raw score;

 Statistics for Psychology, Sixth Edition

 ALWAYS LEARNING

 Arthur Aron | Elliot J. Coups | Elaine N. Aron

 Copyright © 2013 by Pearson Education, Inc. All Rights Reserved



 Suppose the people living in a particular city have a mean score of 40 and a SD of 5 on a measure of concern about the environment. Assume that these concern scores are normally distributed. Using the 50%-34%-14% figures, approximately what percentage of people in this city have score a) above 40, b) above 45, c) above 30, d)above 35, e) below 40, f) below 45, g) below 30, h) below 35?



- (a) 50%
- (b) 16%
- (c) 98%
- (d) 84%
- (e) 50%
- (f) 84%
- (g) 2%
- (h) 16%



 The measure of test anxiety has a mean of 15 and a SD of 5. Using a normal curve table, what percentage of student have scores a) above 16, b) above 17, c) above 18, d) below 18, e) below 14?



- (a) 42.07%
- (b) 34.46%
- (c) 27.43%
- (d) 72.57%
- (e) 42.07%



- If a you are sick and if the doctor thinks that you may have a virus in your blood; the doctor wants you to give blood for some tests
- The blood you give is only a small amount of your all circulation system
- However, this small amount of blood can tell many things about the quality all blood amount in your body
- This small amount is called as sample, whole amount of blood is the population
- Another example is when you cook
 - To test whether the food is ready, you take a small sample/ a spoon full of the food
 - This sample tells you the readiness of the whole food (population)

PEARSON

 <u>https://www.mathsisfun.com/data/stan</u> <u>dard-normal-distribution-table.html</u>



- In psychology research, we typically study samples of individuals to make inferences about some larger group (a population).
- For example
 - A sample might consist of the scores of 50 Canadian women who participate in a particular experiment,
 - whereas the population might be intended to be the scores of all Canadian women



- If you want to know something about a population, your results would be most accurate if you could study the entire population rather than a subgroup from it.
- However, in most research situations this is not practical.
- More important, the whole point of research usually is to be able to make generalizations or predictions about events beyond your reach.



- The strategy in almost all psychology research is to study a sample of individuals who are believed to be representative of the general population (or of some particular population of interest).
- More realistically, researchers try to study people who do not differ from the general population in any systematic way that should matter for that topic of research.
- The sample is what is studied, and the population is an unknown about which researchers draw conclusions based on the sample.



- Population
- Sample
- Methods of sampling
 - Random selection
 - Haphazard selection



Methods of sampling

Random selection

- the ideal method of picking out a sample to study is called random selection.
- The researcher starts with a complete list of the population and randomly selects some of them to study.
- An example of random selection is to put each name on a table tennis ball, put all the balls into a big hopper, shake it up, and have a blindfolded person select as many as are needed.
- In practice, most researchers use a computergenerated list of random numbers



- Methods of sampling
 - Haphazard selection
- just taking whoever is available or happens to be first on a list.
- When using haphazard selection, it is surprisingly easy to pick accidentally a group of people that is very different from the population as a whole.



Haphazard selection

- Consider a survey of attitudes about your statistics instructor
- Suppose you give your questionnaire only to other students sitting near you in class.
- Such a survey would be affected by all the things that influence where students choose to sit, some of which have to do with just what your survey is about—how much students like the instructor or the class. Thus, asking students who sit near you would likely result in opinions more like your own than a truly random sample would..

PEARSON

- Unfortunately, it is often impractical or impossible to study a truly random sample. ☺
- Much of the time, in fact, studies are conducted with whoever is willing or available to be a research participant. (voluntariness)
- At best, as noted, a researcher tries to study a sample that is not systematically unrepresentative of the population in any known way.



Statistical terminology for samples and populations

- The mean, variance, and standard deviation of a population are called **population parameters**.
- A population parameter usually is *unknown* and can be estimated only from what you know about a sample taken from that population
- Population parameters are usually shown as Greek letters (e.g., μ).
- The symbol for the mean of a population is μ , the Greek letter mu. The symbol for the variance of a population is σ^2 , and the symbol for its standard deviation is σ , the lowercase Greek letter sigma.



Statistical terminology for samples and populations

- The mean, variance, and standard deviation you figure for the scores in a sample are called sample statistics.
- A sample statistic is *figured* from known information.
- Sample statistics are what we have been figuring all along and are expressed with the Roman letters you learned in Chapter 2: M, SD², and SD.
- The population parameter and sample statistic symbols for the mean, variance, and standard deviation are summarized in Table 3-2.



Table 3-2	Population Parameters and Sample Statistics	
	Population Parameter (Usually Unknown)	Sample Statistic (Figured from Known Data)
Basis:	Scores of entire population	Scores of sample only
Symbols:		
Mean	μ	М
Standard deviation	σ	SD
Variance	σ^2	SD ²

Table 3-2 Population Parameters and Sample Statistics

Statistics for Psychology, Sixth Edition ALWAYS LEARNING Arthur Aron | Elliot J. Coups | Elaine N. Aron Copyright © 2013 by Pearson Education, Inc. All Rights Reserved

PEARSON

Probability

- The purpose of most psychological research is to examine the truth of a theory or the effectiveness of a procedure.
- But scientific research of any kind can only make that truth or effectiveness seem more or less likely; it cannot give us the luxury of knowing for certain.
- In particular, probability is very important in inferential statistics, the methods psychologists use to go from results of research studies to conclusions about theories or applied procedures



Interpretations of Probability

- Probability is the expected relative frequency of a particular outcome
- An outcome is the result of an experiment (or just about any situation in which the result is not known in advance, such as a coin coming up heads or it raining tomorrow).
- **Frequency** is how many times something happens.
- The *relative frequency* is the number of times something happens relative to the number of times it could have happened;
 - that is, relative frequency is the proportion of times something happens.



Interpretations of Probability

- Think probability in terms of the likelihood of a particular outcome occurring.
- If the probability is very low, then the outcome is unlikely to occur,
- but if the probability is higher, then it is more likely that the outcome will happen.



Figuring probability

 Probabilities are usually figured as the proportion of successful possible outcomes— the number of possible successful outcomes divided by the number of *all* possible outcomes

 $Probability = \frac{Possible successful outcomes}{All possible outcomes}$

- Eg. There are 10 male sts in our class of 50 sts
- If I were to pick smo from the class at random, the probability of picking a male would be 10/50 = 0.2



Range of probabilities

- A probability (olasılık) is a proportion (oran)
- A proportion cannot be less than 0 or greater than 1
- In terms of percentages, proportions range from 0% to 100%
- When the probability of an event is 0, the event is completely *impossible;* it cannot happen.
- But when the probability of an event is low, say 5% or even 1%, the event is *improbable* or *unlikely*, but not impossible.



Range of probabilities

- Probability is usually symbolized by the letter *p*.
- The actual probability number is usually given as a decimal, though sometimes fractions or percentages are used.
- Example:
 - A 50-50 chance is usually written as p = .5, but it could also be written as $p = \frac{1}{2}$ or p = 50%
- It is also common to see probability written as being less than some number, using the "less than" sign.
 For example, p < .05 means "the probability is less than .05."
- Probability and the normal distribution
 - Normal distribution as a probability distribution



Probability, Z scores and the Normal Distribution

- If you think of probability in terms of the proportion of scores, probability fits in well with frequency distributions
- Example:
- If I were to select a person w/ stress rating of 8 from this group of 30 sts at random, there would be 5 chances out of 30; p = 5/30 = .17

Table 1-3Frequency Tableof Number of Students Rating EachValue of the Stress Scale

Stress Rating	Frequency	Percent
0	1	3.3
1	1	3.3
2	1	3.3
3	2	6.7
4	1	3.3
5	2	6.7
6	4	13.3
7	7	23.3
8	5	16.7
9	3	10.0
10	3	10.0

Source: Data based on Aron et al. (1995).

Probability, Z scores and the Normal Distribution

- You can also think of the normal distribution as a probability distribution
- The percentage of scores between any two Z scores is the same as the probability of selecting a score between those two Z scores.
- Example:
- The probability of a score being between the mean and a Z score of +1 is about .34

• *p* = .34



End of Chapter 3

 Statistics for Psychology, Sixth Edition

 ALWAYS LEARNING

 Arthur Aron | Elliot J. Coups | Elaine N. Aron

 Copyright © 2013 by Pearson Education, Inc. All Rights Reserved

