UNDERSTANDING RESEARCH RESULTS: STATISTICAL INFERENCE

A FEW TERMS

Variable

1 2 3 (4) 5

Statistic

Average = 3.75

Parameter

Average = 3.72

A FEW TERMS

Data analysis

THE GATHERING, DISPLAY, AND SUMMARY OF DATA;

Probability

THE LAWS OF CHANCE, IN AND OUT OF THE CASINO;

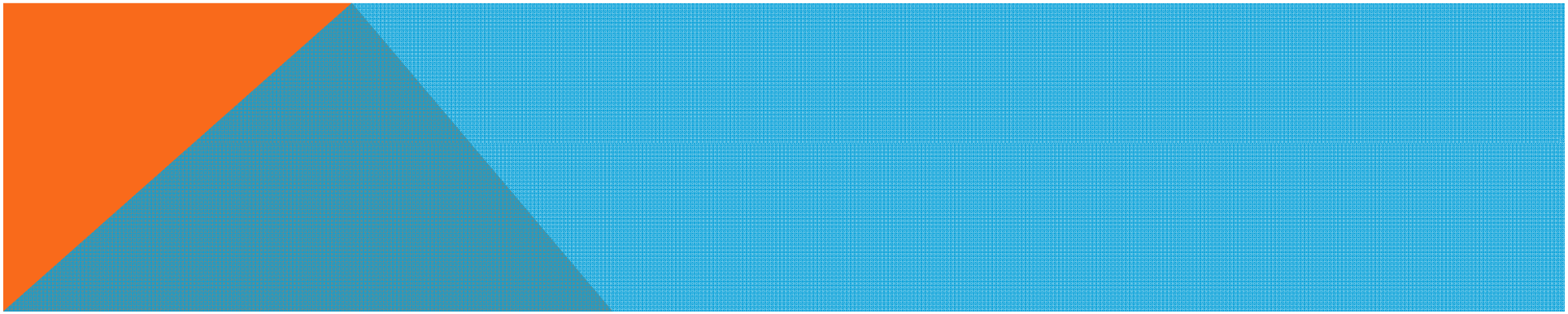
Statistical inference

THE SCIENCE OF DRAWING STATISTICAL CONCLUSIONS FROM SPECIFIC DATA, USING A KNOWLEDGE OF PROBABILITY.

SAMPLES AND POPULATIONS

- ✓ Inferential statistics are necessary because
  - ✓ The results of a given study are based on data obtained from a single *sample* of researcher participants and
  - ✓ Data are not based on an entire population of scores

- ✓ Allows conclusions on the basis of sample data



INFERENTIAL STATISTICS

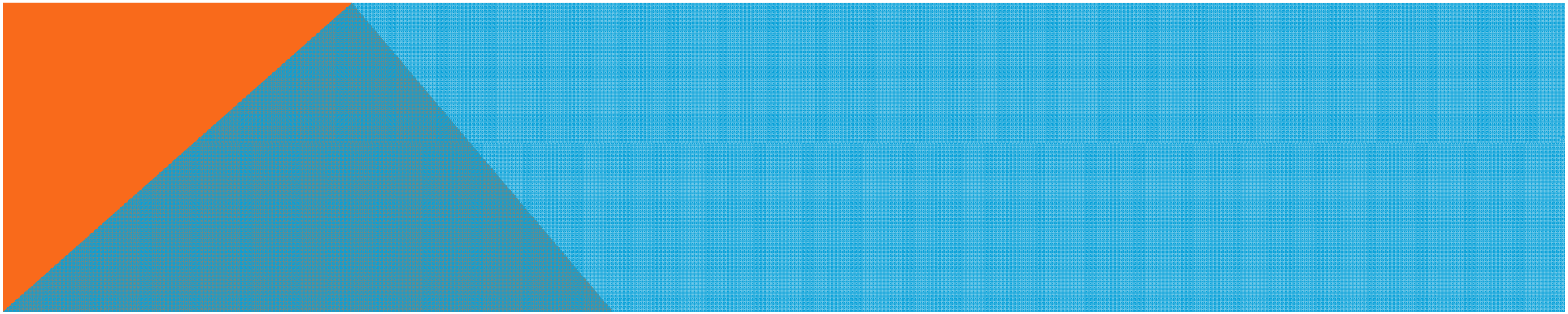
- ✓ Allow researchers to make inferences about the true differences in populations of scores based on a *sample* of data from that population
- ✓ Allows that the difference between sample means reflects *random error* rather than a real difference

NULL AND RESEARCH HYPOTHESES

- ✓ Null Hypothesis
  - ✓ H_0: Population Means are Equal
  - ✓ Any differences between groups are due to chance alone

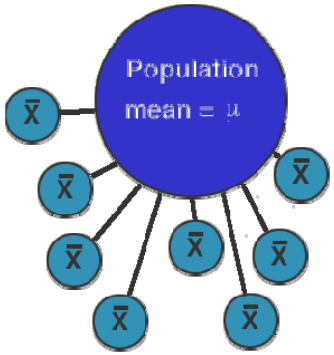
- ✓ Research Hypothesis
  - ✓ H_1: Population Means are Not Equal
  - ✓ Any differences between groups are due to the IV

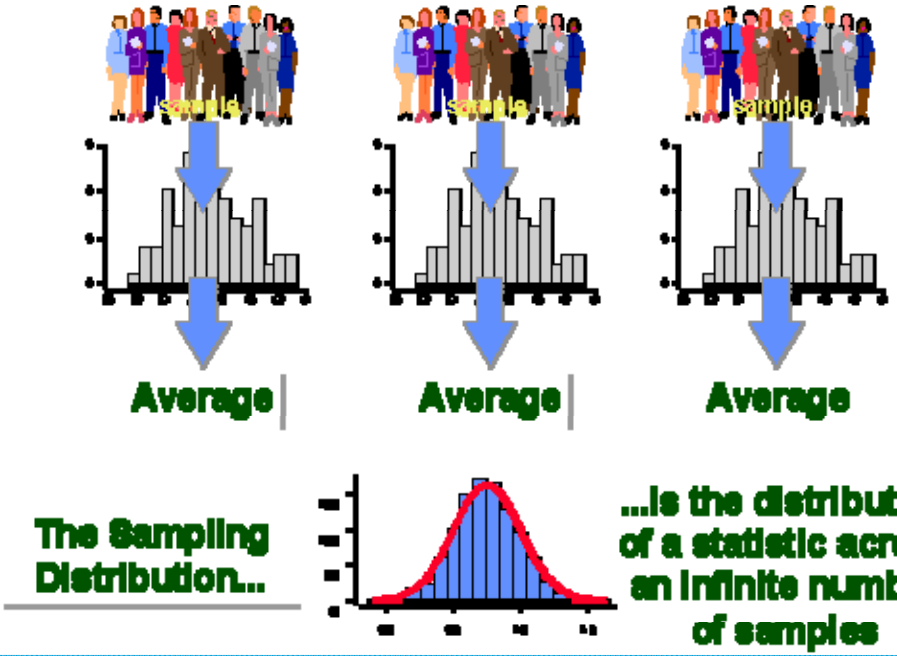
- ✓ Statistical Significance: *Unlikely* to be due to chance alone



PROBABILITY AND SAMPLING DISTRIBUTIONS

- ✓ Probability: The Case of ESP
  - ✓ Are correct answers due to chance or due to something more?
- ✓ Sampling Distributions
  - ✓ Assumes Null is True
  - ✓ Binomial Distributions
- ✓ Sample Size

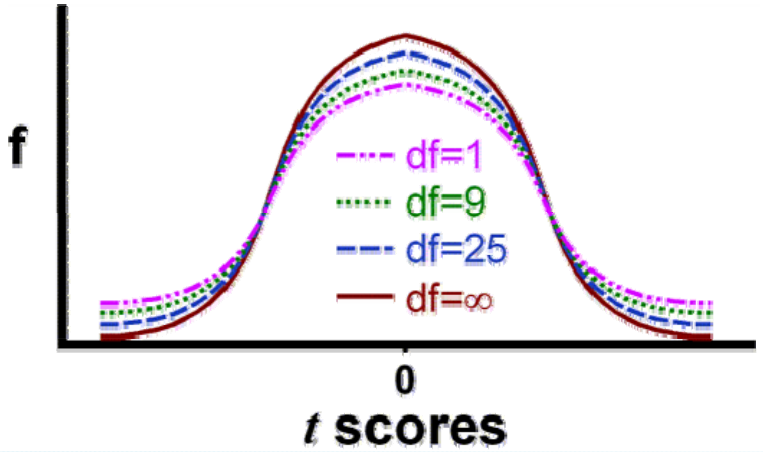




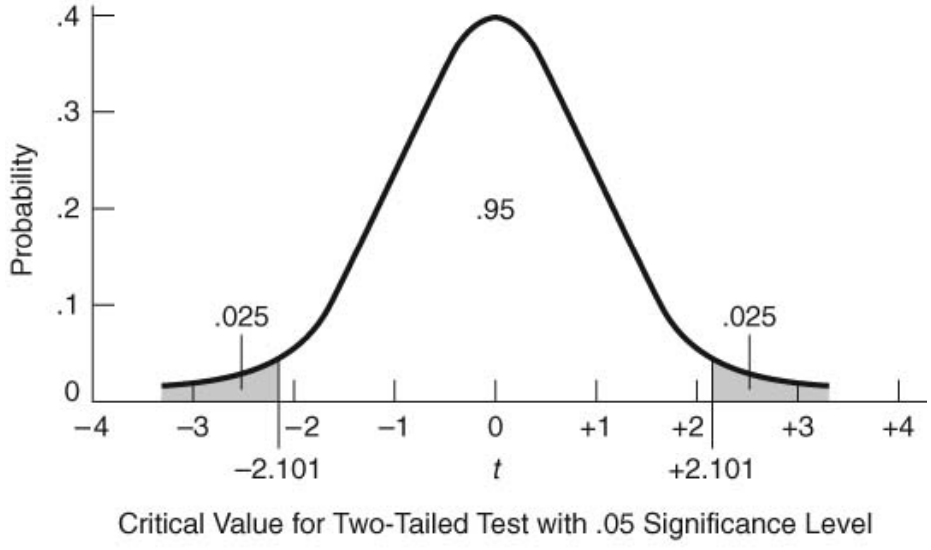
THE *T* TEST

- ✓ *t* value is a ratio of two aspects of the data
  - ✓ The difference between the group means and
  - ✓ The variability within groups

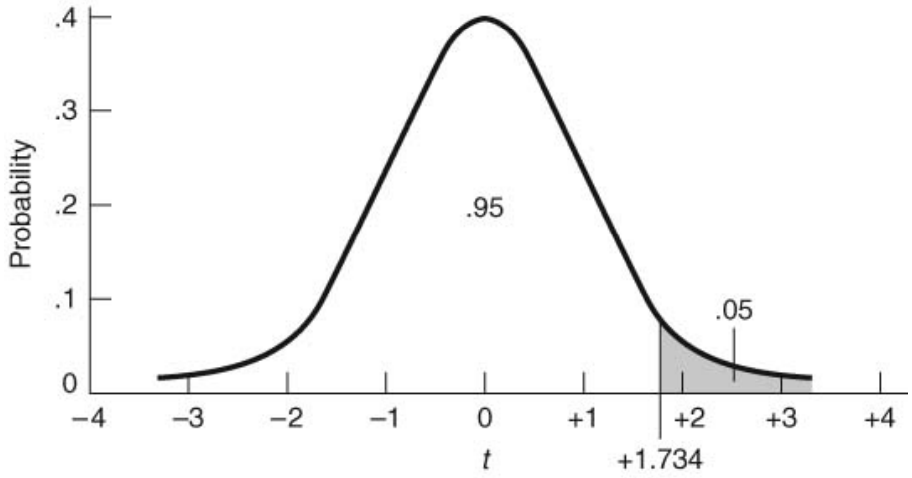
$$t = \frac{\text{between-groups difference}}{\text{within-groups difference}}$$



SAMPLING DISTRIBUTION OF *T* VALUES



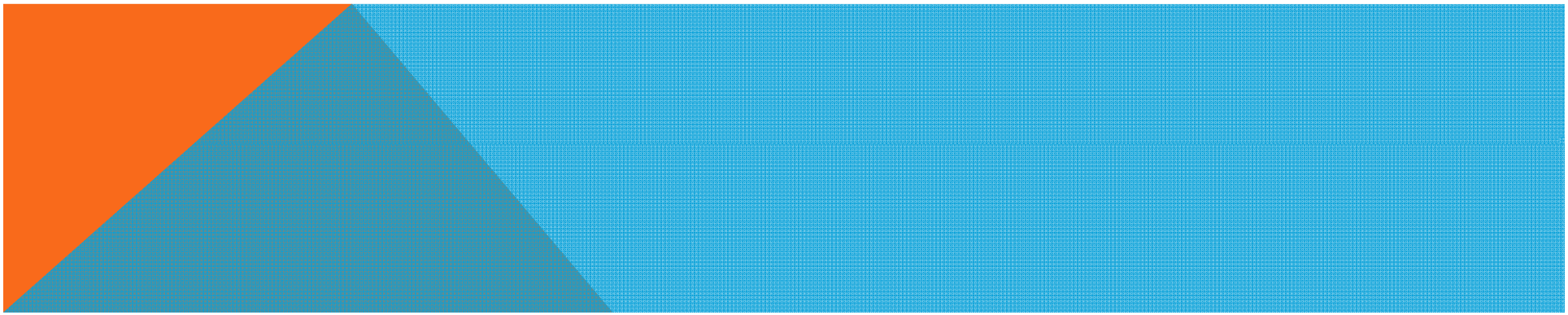
Critical Value for Two-Tailed Test with .05 Significance Level



Critical Value for One-Tailed Test with .05 Significance Level

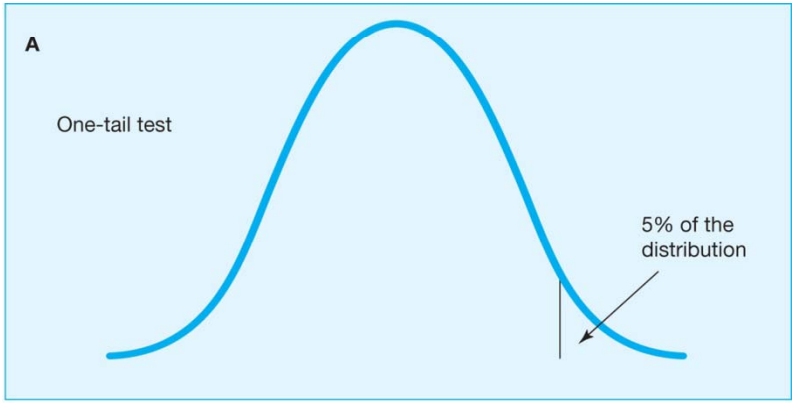
EXAMPLE: THE *T* AND *F* TESTS

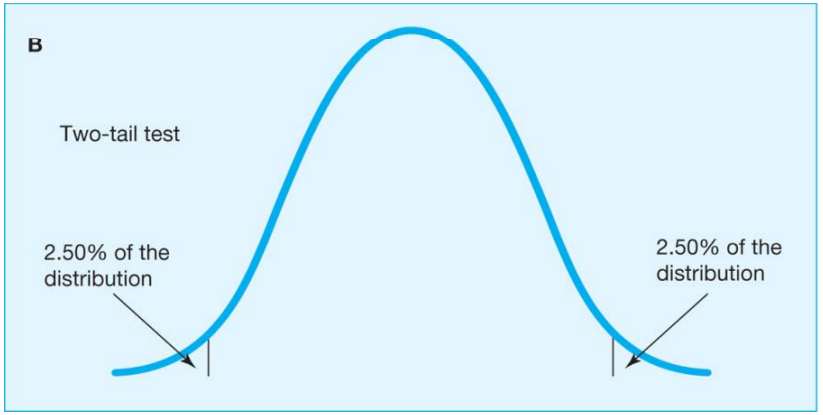
- ✓ Degrees of Freedom
  - ✓ The number of scores that are free to vary when making an estimate.
- ✓ One-Tailed vs. Two-Tailed Tests
  - ✓ Directional vs. Nondirectional Hypotheses?
- ✓ The *F* Test (analysis of variance)
  - ✓ Systematic variance / Between-Groups Variance
  - ✓ Error variance / Within-Groups Variance



ONE-TAILED VS. TWO-TAILED TESTS

What is a directional hypothesis?





Is a one-tailed test valid?

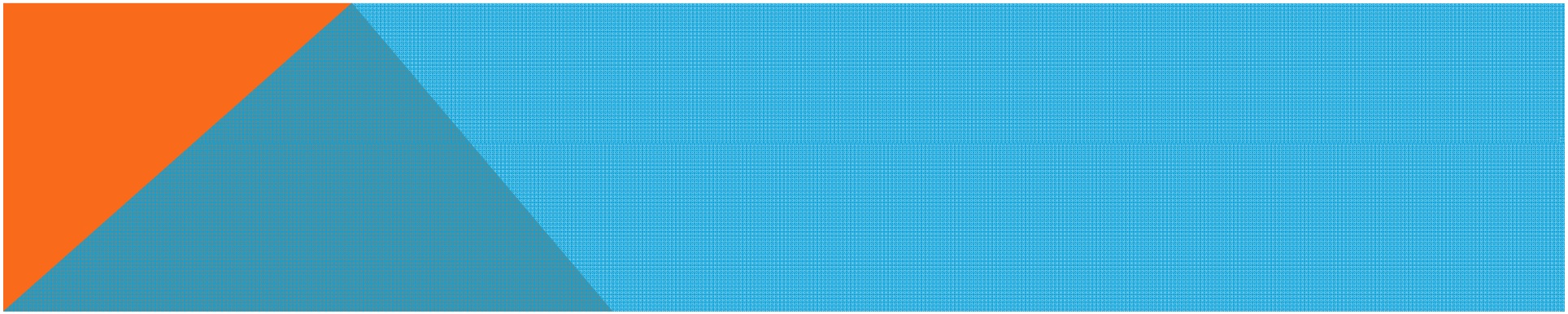
EXAMPLE: THE *T* AND *F* TESTS

- ✓ Calculating Effect Size (*r*, Cohen's *d*, η_p^2 , etc.)
- ✓ Confidence Intervals and Statistical Significance
- ✓ Statistical Significance

TABLE 13.2 Total sample size needed to detect a significant difference for a *t* test

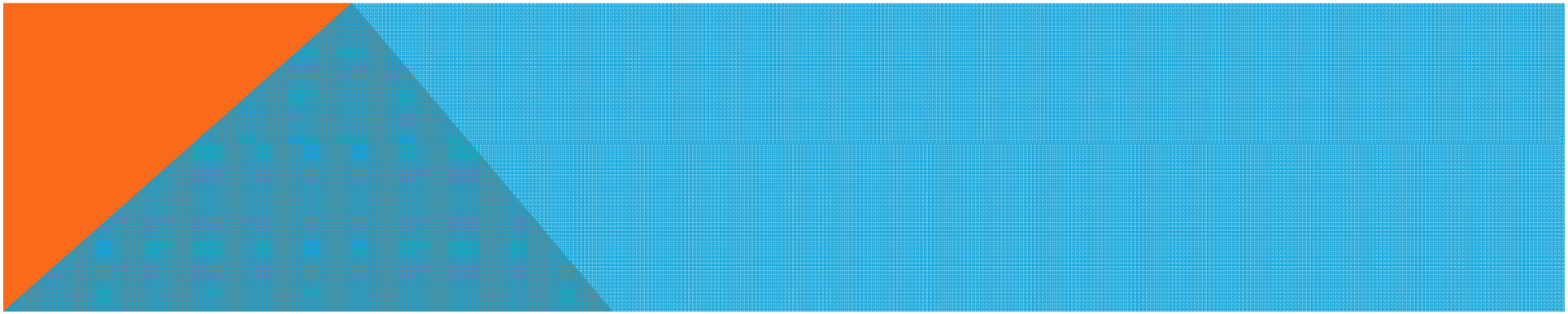
Effect size *r*	Power = .80	Power = .90
.10	786	1052
.20	200	266
.30	88	116
.40	52	68
.50	28	36

Note: Effect sizes are correlations, based on two-tailed tests.

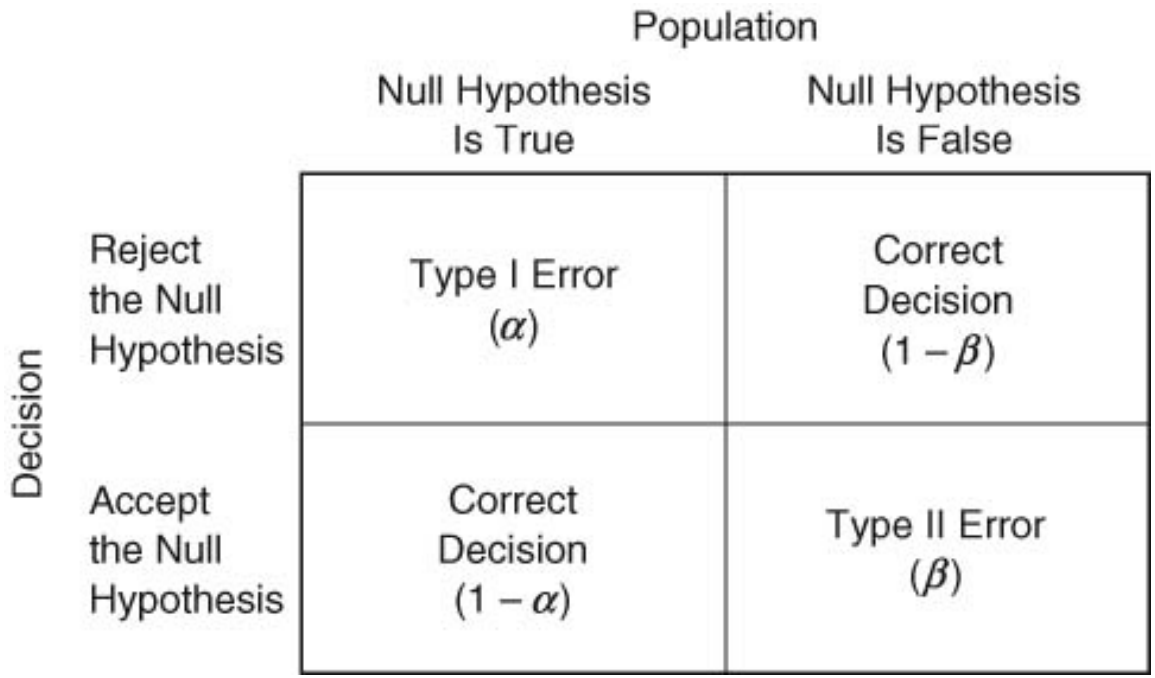


TYPE I AND TYPE II ERRORS

- ✓ Type I Errors
  - ✓ Made when the null hypothesis is rejected but the null hypothesis is actually true
  - ✓ Obtained when a large value of *t* or *F* is obtained *by chance alone*
- ✓ Type II Errors
  - ✓ Made when the null hypothesis is accepted although in the population the research hypothesis is true
  - ✓ Factors related to making a Type II error
    - ✓ Significance (alpha) level
    - ✓ Sample size
    - ✓ Effect size



TYPE I AND TYPE II ERRORS



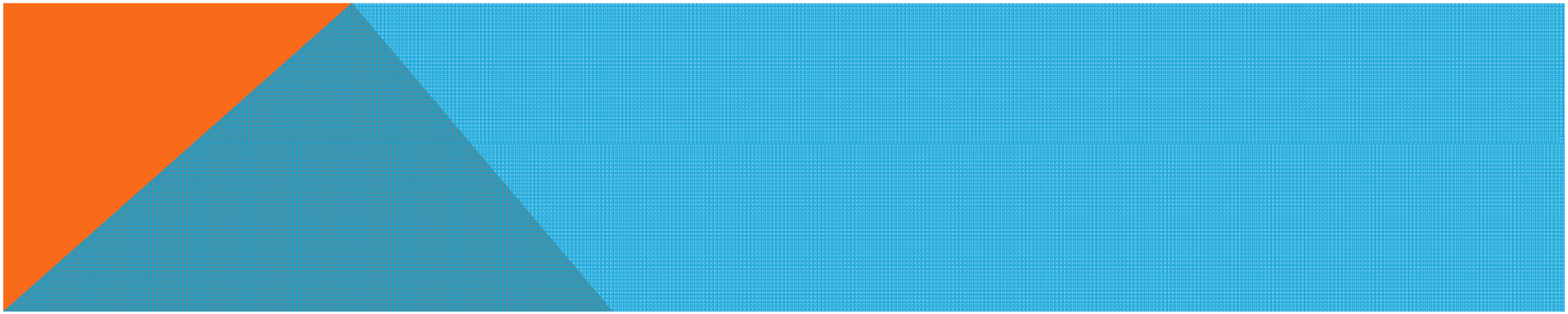
THE EVERYDAY CONTEXT OF TYPE I AND TYPE II ERRORS

Decision	True state: Null is true (innocent)	True state: Null is false (guilty)
Reject null (find guilty)	Type I error	Correct decision
Accept null (find innocent)	Correct decision	Type II error

SIGNIFICANCE LEVEL & STATISTICAL ERRORS

- ✓ Researchers traditionally have used either a .05 or a .01 significance level in the decision to reject the null hypothesis
- ✓ Agreement that the consequences of making a Type I error are more serious than those associated with a Type II error
  - ✓ Arlo's Two Cents: It depends on context.
    - ✓ e.g., A new drug which may cure AIDS

- ✓ Interpreting nonsignificant results
  - ✓ Absence of evidence is not evidence of absence.
  - ✓ Nothing is ever proven or disproven!



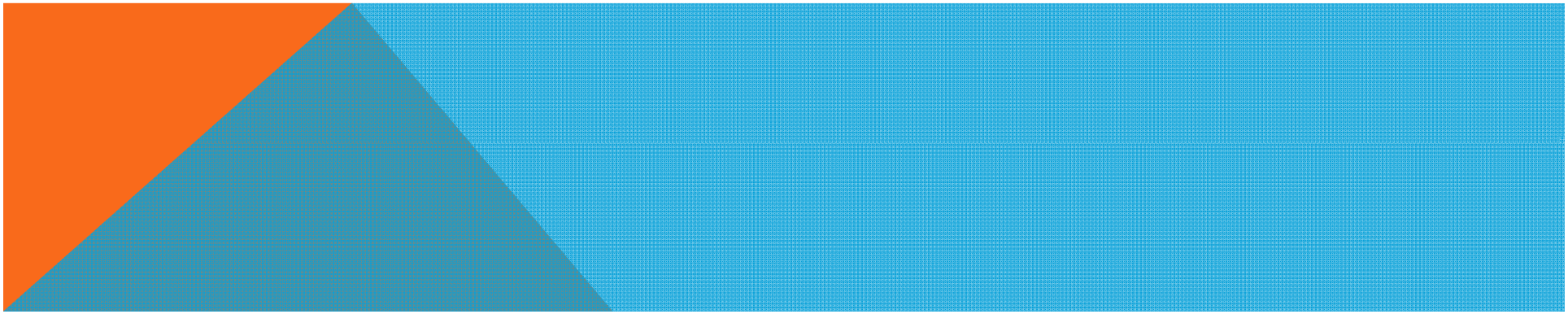
CHOOSING A SAMPLE SIZE: POWER ANALYSIS

Power is a statistical test that determines optimal sample size based on probability of correctly rejecting the null hypothesis

$$\text{Power} = 1 - p(\text{Type II error})$$

Effect sizes range and desired power

- Smaller effect sizes require larger samples to be significant
- Higher desired power demands a greater sample size
- Researchers usually use a power between .70 and .90



IMPORTANCE OF REPLICATIONS

Scientists attach little importance to results of a single study

Detailed understanding requires numerous studies examining same variables

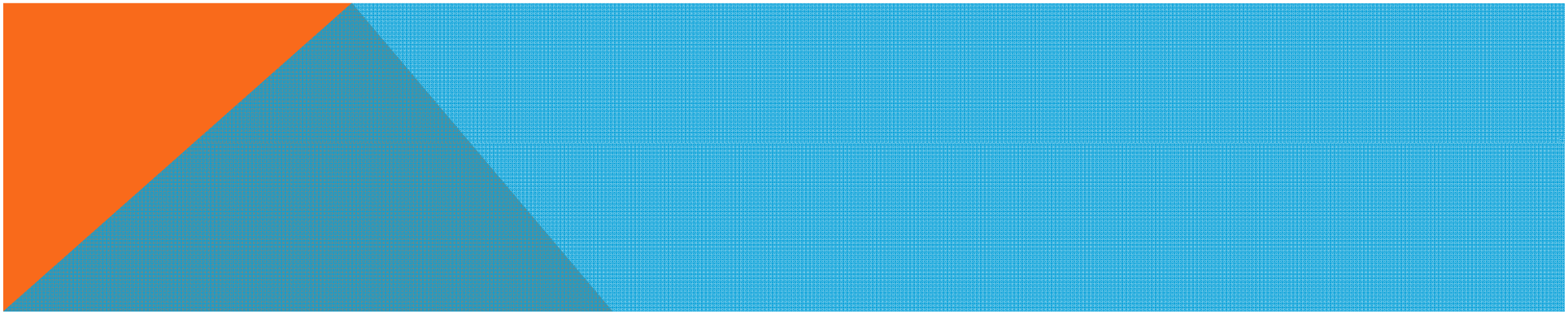
Researchers look at the results of studies that replicate previous investigations

SIGNIFICANE OF PEARSON'S *R* CORRELATION COEFFICIENT

Is the correlation statistically significant?

- Ho: $r = 0$
- H1: $r \neq 0$

It is proper to conduct a t-test to compare the r-value with the null correlation of 0.00 ?



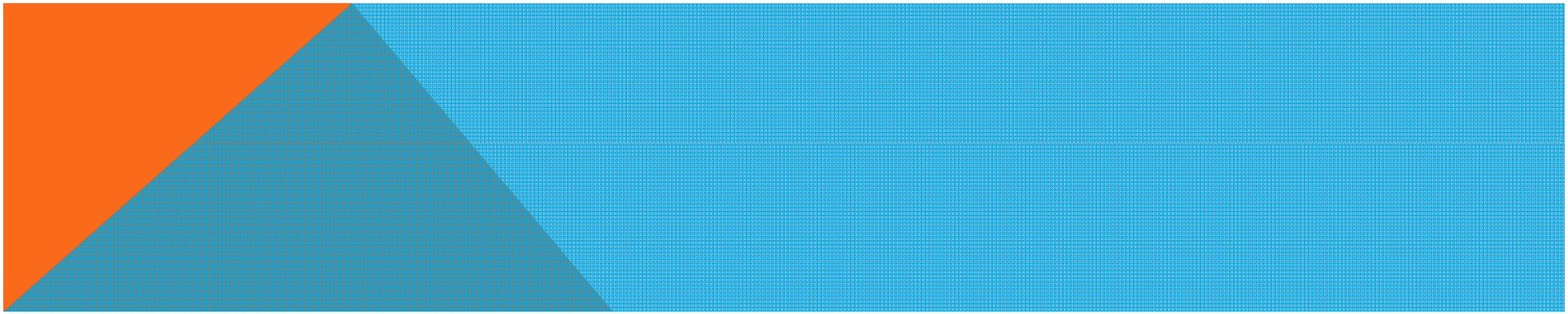
COMPUTER ANALYSIS OF DATA

- ✓ Software Programs include
  - ✓SPSS
  - ✓SAS
  - ✓Minitab
  - ✓Microsoft Excel
- ✓ Steps in analysis

✓Input data
- ✓Rows represent cases or each participant's scores
- ✓Columns represent a participant's score for a specific variable

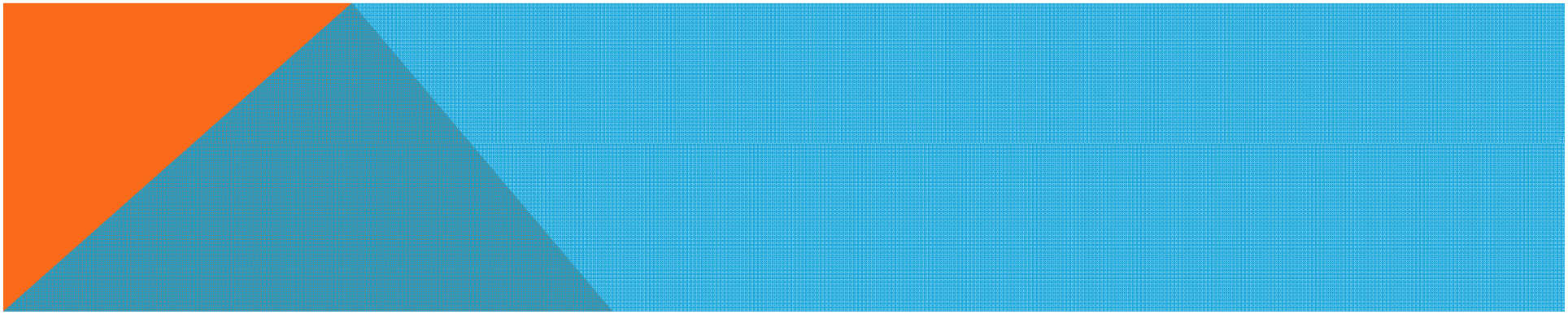
✓Conduct analysis

✓Interpret output



SELECTING THE APPROPRIATE SIGNIFICANCE TEST

IV	DV	Statistical Test
Nominal Male-Female	Nominal Vegetarian – Yes / No	Chi Square
Nominal (2 Groups) Male-Female	Interval / Ratio Grade Point Average	*t* test
Nominal (3 groups) Study time (Low, Medium, High)	Interval / Ratio Test Score	One-way ANOVA
Interval / Ratio Optimism Score	Interval / Ratio Sick Days Last Year	Pearson's correlation



SELECTING THE APPROPRIATE SIGNIFICANCE TEST

- ✓ Multiple Independent Variables
- ✓Nominal Scale Data – ANOVA Factorial Design
- ✓Ordinal Scale Data – no appropriate test is available
- ✓Interval or Ratio Scale Data – Multiple Regression