GETTING STARTED WITH STATISTICS FOR LIBRARIANS

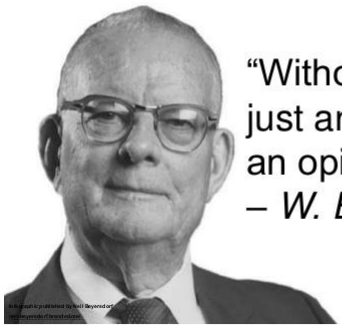
Jin Wu
Norris Medical Library
University of Southern California
jwu5@usc.edu

COURSE OUTLINE

- Examine examples of research papers
- Define key statistical concepts
- Use examples and exercise to verify understanding
- Re-examine the research papers to reinforce the concept.

LEARNING OUTCOMES

- Understand basic statistical terms such as standard deviation, t-test, p value, etc.
- Identify test measures used in published research studies
- Interpret the data, graphs, and statistical output used in the research findings



"Without data you're just another person with an opinion."
– *W. Edwards Deming*

EXAMINE EXAMPLES OF RESEARCH PAPERS

Citation patterns of online and print journals in the digital age

Sandra L. De Groote, MLIS, AHIP

See end of article for author's affiliation.

DOI: 10.3163/1536-5050.96.4.012

increased in 2000 and 2001 compared to 1996, although the numbers slightly decreased in 2004 and 2005. Interestingly, journals available only in print were cited on average more than the journals available in print and online formats.

A repeated measures analysis of variance (ANOVA) test was performed to examine the effects of online journals on the citation patterns of urban authors. The analysis did not find a statistically significant interaction between year of citation and print status. In other words, the findings suggested that over time, journals in print were not less likely to be cited and journals available online were not more likely to be cited, when compared to journals cited prior to the introduction of online journals. An additional ANOVA test was run on a subset (journals in the MEDLINE Abridged Index Medicus [AIM] subset) of the above journals with similar results (Table 2).

Regional campus. A total of 760 journals were cited by authors at this campus, of which 564 journals cited were excluded from the statistical analysis because

journals remained the most cited during all the years studied. There was a statistically significant interaction between the year of citation and the print status ($F=5.256$, $P<0.001$). In general, the number of cited references decreased during the study years for the print-only journals, while the cited references of the online-and-print journals increased. Cited references for the online-only journals also increased over the study period, with the exception of a decrease in 2005. Of note was the decrease overall in the number of articles cited in 2005. As also noted at the urban setting, journals available in print were cited on average more than the other journals.

An ANOVA test of the AIM journals also showed an increase in citing the online-and-print journals and a decrease in citing the print-only journals ($F=2.194$, $P<0.038$). These findings suggested that online journals have had an impact on the citation patterns of authors on the regional campus, where only a small print journal collection was available. Researchers were citing the journals available online more and citing the journals available only in print less.

De Groote, (2008) p.365

Clinical and academic use of electronic and print books: the Health Sciences Library System e-book study at the University of Pittsburgh

Barbara L. Folb, MM, MLS, MPH; Charles B. Wessel, MLS; Leslie J. Czechowski, MA, MLS

See end of article for authors' affiliations.

DOI: 10.3163/1536-5050.99.3.009

Table 3
Reported e-book use by role at University of Pittsburgh Medical Center (UPMC) or University of Pittsburgh

Affiliation and role*	Reported use of e-books n	(%)
UPMC (n=435)†		
Intern, resident, or fellow (n=91)	73	(80.2%)
Attending physician (n=71)	52	(73.2%)
Researcher (n=74)	42	(56.8%)
Other (n=25)	14	(56.0%)
Other patient care (n=35)	19	(54.3%)
Support staff (n=43)	18	(41.9%)
Nurse (n=68)	28	(41.2%)
Administrator (n=28)	8	(28.6%)
University of Pittsburgh (n=648)‡		
Postdoctoral or fellow (n=74)	54	(73.0%)
Faculty or teaching role (n=215)	139	(64.7%)
Graduate or medical student (n=205)	127	(62.0%)
Staff (n=84)	43	(51.2%)
Undergraduate (n=45)	22	(48.9%)
Other (n=25)	12	(48.0%)

* Respondents can appear in more than one category. Respondents with UPMC email addresses indicating roles at the university are included in both categories.
† χ^2=48.051, *df*=7, *P*=0.000.
‡ χ^2=13.705, *df*=5, *P*=0.018

Barriers and facilitators to e-book use

Respondent awareness and use of the e-book collection. Most respondents (n=599/914, 65.5%) recalled seeing information about e-books on the HSLS website, although slightly fewer (n=505/911, 55.4%) reported using an HSLS e-book. Use of e-books to look up brief factual information was reported by 56.6% (n=516/911), while use for in-depth study was reported by 41.9% (n=383/913).

Use and rating of e-book search tools. The utility of the 5 HSLS e-book search tools, Google Books, and the Amazon Search Within the Book feature was rated by 863 respondents, as summarized in Figure 1. The federated full-text search tool was used by the largest percent of respondents (n=580/863, 67.2%) and was rated moderately to extremely useful by 74.3% (n=431/580) who used it. Google Books was also rated as moderately to extremely useful by 74.3% (n=373/502) who used it. They gave the lowest ratings to the library catalog (PITTCat), with 61.2% (n=306/500) rating it moderately to extremely useful.

Folb, (2011) p.222

The Librarian Leading the Machine: A Reassessment of Library Instruction Methods

Katie Greer, Amanda Nichols Hess, and Elizabeth W. Kraemer

TABLE 2
Descriptive Statistics

Type	N	Mean	Std. Deviation	Std. Error Mean
Experimental Group	129	20.12	3.798	.334
Control Group	128	20.20	3.775	.334

Results

Before analysis commenced, data from students who had indicated they were underage or did not wish to have their data included in the study were removed. The remaining, anonymized data resulted in a near-equal sample size for each instructional method: 128 from the blended instruction control group and 129 from the online-only group. The authors used SPSS software to analyze the data. Table 2 provides the descriptive statistics and table 3 provides the results of an independent samples t-test.

In regard to the research question "Does the delivery format of information literacy instruction in face-to-face course sections affect attainment of student learning outcomes?" the data suggest, in this case, that it does not. The means of the two groups are nearly identical, as shown in table 2. An independent t-test, which is a generally accepted statistical measure of difference, confirmed that the groups did not differ significantly ($p = 0.88$; $p > 0.05$ confirms that there is no significant difference).

TABLE 3
Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal Variances Assumed	.835	.362	–.151	255	.880	–.07	.472	–1.002	.859
Equal Variances Not Assumed			–.151	254.999	.880	–.07	.472	–1.002	.859

Greer, (2009) p.293-294

Trends in health sciences library and information science research: an analysis of research publications in the *Bulletin of the Medical Library Association* and *Journal of the Medical Library Association* from 1991 to 2007*

Sally A. Gore, MS, MS LIS; Judith M. Nordberg, MLIS; Lisa A. Palmer, MSLS, AHIP; Mary E. Piorun, MSLS, MBA, AHIP

See end of article for authors' affiliations.

DOI: 10.3163/1536-5050.97.3.009

None	344	(72.6)	
Government	66	(13.9)	
Association	30	(6.3)	
Own institution	21	(4.4)	
Other	13	(2.7)	
Total number of authors†			=0.0019*
1	180	(38.0)	
2	142	(30.0)	
3	83	(17.5)	
4	33	(7.0)	
5	18	(3.8)	
6+	18	(3.8)	
Total number of pages‡			=0.0055*
1–4	89	(18.8)	
5–9	310	(65.4)	
10–14	68	(14.3)	
15–19	4	(0.8)	
20+	3	(0.6)	
Total number of citations§			<0.0001*
1–4	25	(5.3)	
5–9	87	(18.4)	
10–14	104	(21.9)	
15–19	91	(19.2)	
20+	167	(35.2)	

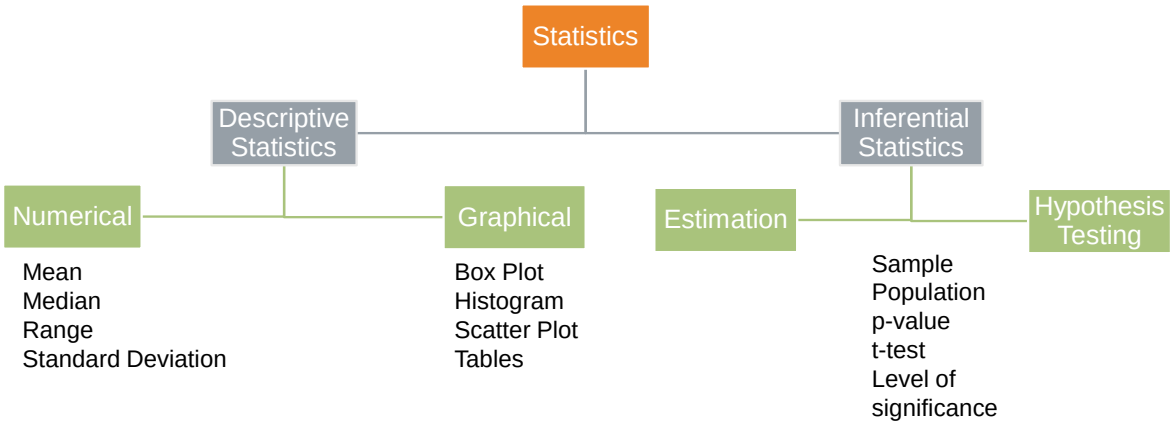
* Probability value for comparison of these results with those reported by Dimitroff [6]. $P^1 \leq 0.05$ is statistically significant.
† Mean: 2.2, median: 2, SD: 1.3.
‡ Mean: 7.0, median: 7, SD: 3.2.
§ Mean: 19.1, median: 16, SD: 15.0.

Gore, (2009) p.206

BASIC STATISTICAL CONCEPTS

- **What is statistics?**
  - Statistics is the study of the collection, analysis, interpretation, preparation, and organization of data.[5]
- **Why is it important?**
  - Explain what happens
  - Evaluate the credibility and usefulness of information
  - Make sound decisions

TYPES OF STATISTICS



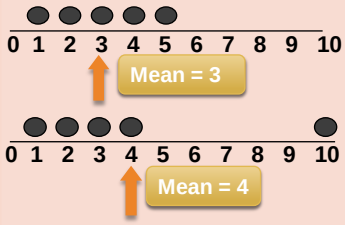
DESCRIPTIVE STATISTICS

- Numbers used to describe the population data
- They do not actually test any hypotheses (or yield any p-values)
- Types:
  - Measures of Central Tendency
    - Mean
    - Median
    - Mode
  - Measures of Dispersion
    - Range
    - Quartile
    - Standard Deviation
  - Measures of Frequency
    - Histogram
    - Bar Chart
- Limitations:
  - Cannot use the data collected to generalize to other people or objects
  - Don't always need to generalized to other populations

MEASURES OF CENTRAL TENDENCY

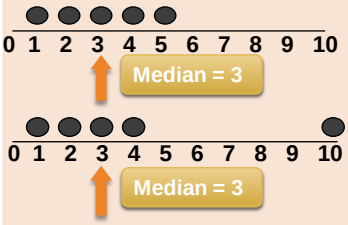
Mean

- The most common measure of central tendency
- Average
- Affected by extreme values



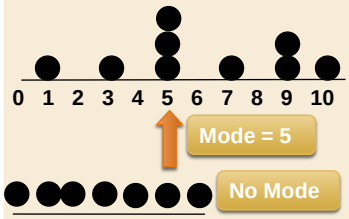
Median

- Middle
- Not affected by extreme values



Mode

- Most frequent value
- Not affected by extreme values
- There may be no mode or several modes



MEASURES OF DISPERSION

Range

- The difference between the highest and lowest score in a data set.

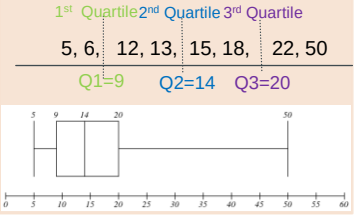
Example: 2, 5, 16, 35, 36, 40, 55

Range=55-2=3

Quartiles

- Quartiles tell us about the spread of a data set by breaking the rank-ordered data set into quarters.

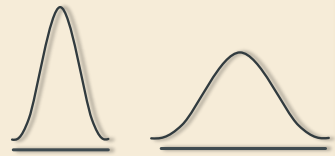
Example:



Standard Deviation

- Shows how much variation there is from the average.

Example:



Small standard deviation

Large standard deviation

EXAMPLE

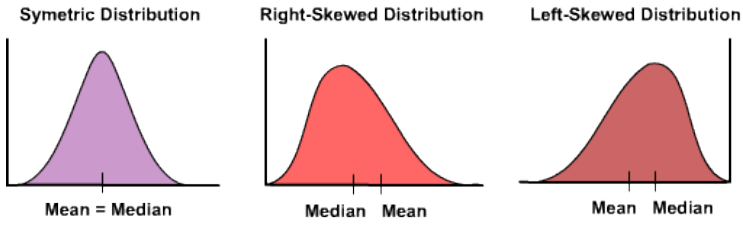
Table 1
Descriptive statistics for *Bulletin of the Medical Library Association (BMLA)* and *Journal of the Medical Library Association (JMLA)* research articles, 1991–2007

Variable	Number of research articles (n=474)	(%)	Adjusted probability (P^1) value*
Total number of authors†			=0.0019*
1	180	(38.0)	
2	142	(30.0)	
3	83	(17.5)	
4	33	(7.0)	
5	18	(3.8)	
6+	18	(3.8)	
Total number of pages‡			=0.0055*
1–4	89	(18.8)	
5–9	310	(65.4)	
10–14	68	(14.3)	
15–19	4	(0.8)	
20+	3	(0.6)	
Total number of citations§			<0.0001*
1–4	25	(5.3)	
5–9	87	(18.4)	
10–14	104	(21.9)	

* Probability value for comparison of these results with those reported by Dimitroff [6]. $P^1 \leq 0.05$ is statistically significant.
† Mean: 2.2, median: 2, SD: 1.3.
‡ Mean: 7.0, median: 7, SD: 3.2.
§ Mean: 19.1, median: 16, SD: 15.0.

Gore, (2009) p.206

WHAT DO THESE NUMBERS TELL US?



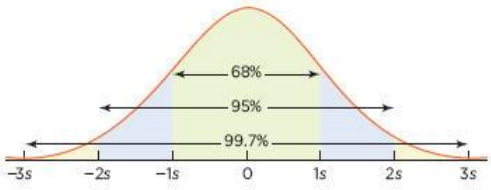
† Mean: 2.2, median: 2, SD: 1.3.
‡ Mean: 7.0, median: 7, SD: 3.2.
§ Mean: 19.1, median: 16, SD: 15.0.

Mean>Median, data is slightly right skewed CV=1.3/2.2=0.65
Mean=Median, data is symetrically distributed CV=3.2/7=0.46
Mean>Median, data is slightly right skewed CV=15/19.1=0.78

CV (Coefficient of Variance)=Standard Deviation/Mean
Distributions with CV<1 are considered low-variance. Lower variation, data is more reliable.
Distributions with CV>1 are considered high-variance. Greater variation, data is less reliable.

EMPIRICAL RULE

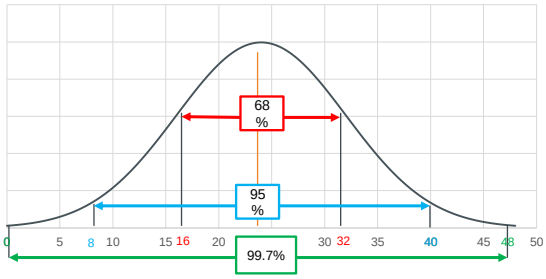
Interval	Percentage of Data
$\bar{y} - s$ to $\bar{y} + s$	68
$\bar{y} - 2s$ to $\bar{y} + 2s$	95
$\bar{y} - 3s$ to $\bar{y} + 3s$	99.7



Copyright © 2011 Pearson Education, Inc.

If mean=24, standard deviation=8.

Interval	Percentage of Data
24-8=16 to 24+8=32	68%
24-2*8=8 to 24+2*8=40	95%
24-3*8=0 to 24+3*8=48	99.7%



EXERCISE

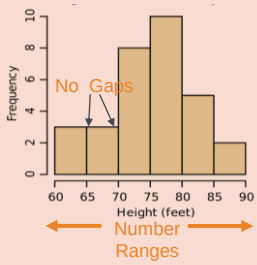
	Mean	Median	Standard Deviation
Data Set A	12	15	6
Data Set B	12	12	6
Data Set C	12	9	15
Data Set D	12	13	3

Questions:
1. Which data set is right-skewed?
2. Which data set is more dispersed?
3. Which data set is more clustered?

MEASURES OF FREQUENCY

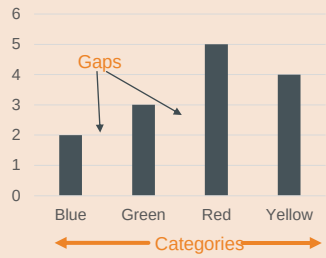
Histogram

- A chart that displays a summary of the frequency distribution of classes that fall within certain lower and upper limits in a set of data.



Bar Charts

- A graphical display of data using bars of different heights.



EXAMPLE

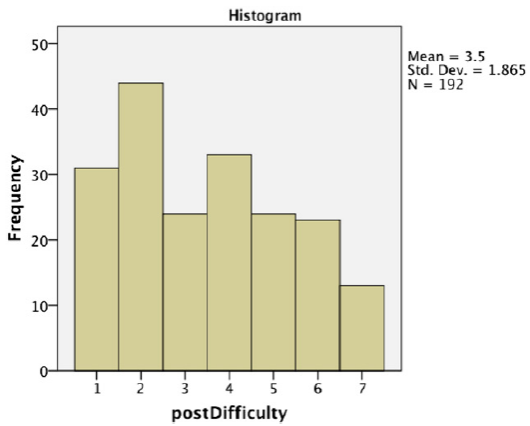
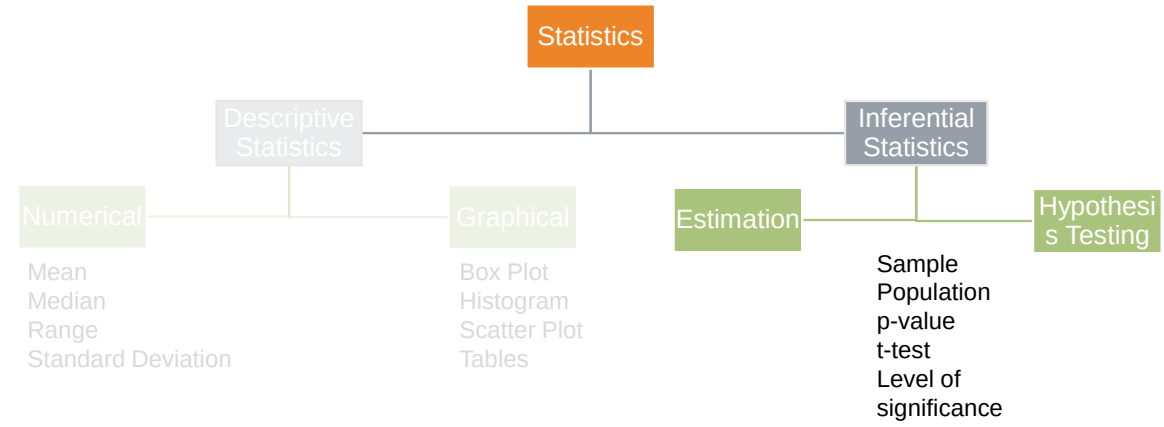


Fig. 1. Distribution of task difficulty ratings in all task sessions.

Liu (2015) p.332

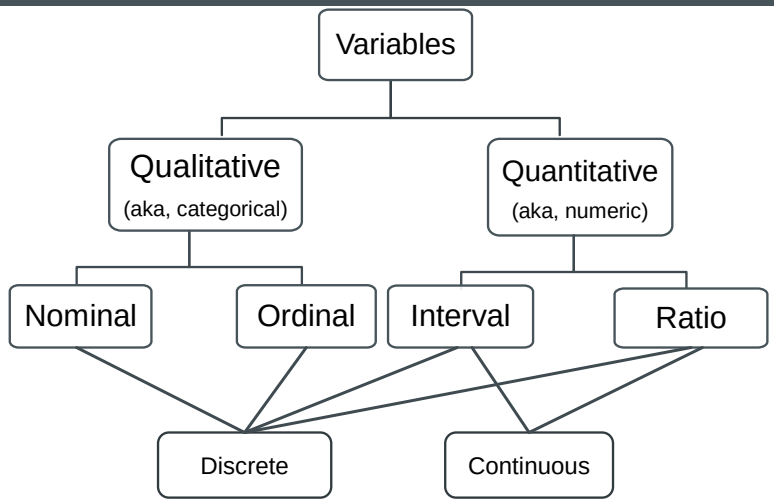
TYPES OF STATISTICS



INFERENTIAL STATISTICS

- Did the event happen by chance or by underlying cause?
- Methods of inferential statistics:
  - Estimation of parameters
  - Testing of statistical hypotheses
- Conclusions are never 100% certain
  - Uncertainty is introduced by random error
  - We can understand how confident we are through confidence interval
- Limitations:
  - Due to random error, there's a small probability that your conclusion might not be right.
  - It doesn't give you causation. It only gives relationship.

VARIABLES



SCALES OF DATA (NOIR)

Nominal

- Counts by category
- Cannot be quantified
- Cannot be assigned any order

Examples
- Gender (Male, Female, Transgender)
- Eye Color (Blue, Green, Black)
- Marital Status (Married, Single, Divorced)

Ordinal

- Counts by category
- Cannot be quantified
- Can be ranked logically

Examples
- Age (Young, Middle-aged, Old)
- Satisfaction (Unsatisfied, Satisfied, Very satisfied)
- Frequency(Never, Sometimes, Often, Always)

Interval

- Difference between two values is meaningful
- No baseline

Examples
- Temperature(-20°, 0°, 20°)

Ratio

- Difference between two values is meaningful
- With a true zero.

Examples
- Height (5'5, 5'8, 6'4)
- Weight (120lbs, 150lbs, 180lbs)
- Income ($0, $50, $100)

SUMMARY OF SCALES OF DATA

Provides	Nominal	Ordinal	Interval	Ratio
The “order” of value is known		✓	✓	✓
Mode	✓	✓	✓	✓
Median		✓	✓	✓
Mean			✓	✓
Can quantify the difference between each value			✓	✓
Can add or subtract value			✓	✓
Can multiple and divide values				✓
Has “true zero”				✓

EXERCISE

What are the **types of variables** measured by the following survey questions?

WHY DOES IT MATTER?

- Different types of data allow for different types of data analysis

	Nominal	Ordinal	Interval/Ratio
Nominal	Phi (2X2 only); Lambda; Cramer's V; Chi Square; Fisher's Exact	Lambda; Cramer's V; **Chi Square**; Fisher's Exact	
Ordinal		Gamma; Kendall's Tau-b; Spearman's Rho; Yule's Q (2X2 only)	Kendall's Tau-b; Spearman's Rho
Interval/Ratio	**t Test** & Mann-Whitney *U* (two attributes of the nominal/ordinal variable); ANOVA (three or more attributes of the nominal/ordinal variable)		Pearson's r; **Regression analysis**

Luo, (2016) slides 55

COMMON INFERENTIAL STATISTICS TEST MEASURES

Test of Relationships	Test of Group Differences	Test of Repeated Measures	Tests Using Categorical Data
Pearson r correlation	Independent t-test	Dependent t-test	Chi-Square Test of Independence
Linear/Multiple Regression	ANOVA	Repeated Measures ANOVA	Logistic Regression

To decide which test to use, check out: https://cyfar.org/types-statistical-tests or consult with a statistician at your institution.

KEY ELEMENTS OF SIGNIFICANCE TESTING

- Null Hypothesis
- Measure of Central Tendency
- Standard Deviation
- Risk of Being Wrong (alpha) 0.05 or 0.25 or 0.01 or 0.001

FIVE STEPS IN HYPOTHESIS TESTING

1. Making assumptions
2. Stating the research and null hypotheses and selecting alpha
3. Selecting the sampling distribution and choose which test to use
4. Computing the test statistic
5. Interpreting the results and making a conclusion

EXAMPLE

1. • Making assumptions
2. • Stating the research and null hypotheses and selecting alpha
3. • Selecting the sampling distribution and choose which test to use
4. • Computing the test statistic
5. • Interpreting the results and making a conclusion

Random sample.

Data is normally distributed.

EXAMPLE

1. • Making assumptions
2. • Stating the research and null hypotheses and selecting alpha
3. • Selecting the sampling distribution and choose which test to use
4. • Computing the test statistic
5. • Interpreting the results and making a conclusion

Null Hypotheses:
There's no difference in the learning outcomes between online only and hybrid instruction (both online and face-to-face)

Alpha=0.05

EXAMPLE

1. Making assumptions
2. Stating the research and null hypotheses and selecting alpha
3. Selecting the sampling distribution and choose which test to use
4. Computing the test statistic
5. Interpreting the results and making a conclusion

Experimental group: 129 students who took the online only instruction.
Control group: 128 students who took the blended instruction.

Independent t-test.

EXAMPLE

1. Making assumptions
2. Stating the research and null hypotheses and selecting alpha
3. Selecting the sampling distribution and choose which test to use
4. Computing the test statistic
5. Interpreting the results and making a conclusion

Use statistical software such as SPSS, SAS, etc to calculate the p value and compare it with the alpha researcher selected.

Different test measures generate different values that correspond to different p values at a certain degree of freedom.

Test Measure	Test value
t-Test	t score
ANOVA	F score
Chi-Square	χ^2 score

EXAMPLE

1. Making assumptions
2. Stating the research and null hypotheses and selecting alpha
3. Selecting the sampling distribution and choose which test to use
4. Computing the test statistic
5. Interpreting the results and making a conclusion

If p>0.05, fail to reject the null hypothesis.

- There's no significant difference between the two groups.

If p<0.05, reject the null hypothesis.

- There's a significant difference between the two groups.

5 Interpreting the results

SEM=SD/√N SEM=3.798/√129=0.334
How precisely the mean of the sample estimates the population mean

TABLE 2
Descriptive Statistics

Type	N	Mean	Std. Deviation	Std. Error Mean
Experimental Group	129	20.12	3.798	.334
Control Group	128	20.20	3.775	.334

There is a 95% chance that the range between -1.002 and 0.859 contains the true difference of the mean

Determines if the two groups have about the same or different amounts of variability between two data sets.

P value=.362>0.05, no statistically difference between the variability of the two groups.

TABLE 3
Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal Variances Assumed	.835	.362 = P	-.151	255	.880 =P	-.07	.472	-1.002	.859
Equal Variances Not Assumed			-.151	254.999	.880	-.07	.472	-1.002	.859

P value =.880>0.05, no statistically significant difference between the experimental group and the control group

Greer, (2009) p.293-294

5 Making a conclusion

Since **p=0.88>0.005**, so we can not confirm there's a significant difference between online-only instruction and blended instruction.

Conclusion from researcher:
"Although some may wish to interpret these study data as an excuse to move all instruction online and return to simpler times, when librarians did not spend a good portion of their lives on instruction, the authors would be remiss to encourage anyone to immediately do so. The outcomes of this study do not suggest that the computer is mightier than the librarian; rather, the success of the online group as compared to the hybrid group only indicates that carefully crafted online learning objects, which conform to the most recent scholarship of teaching and learning and are responsive to the needs of a specific audience, can be as effective in empowering students to achieve desired learning outcomes as in-classroom instruction. In either case, the librarian plays a pivotal role as the instructional designer. "

Greer, (2009) p.297

EXERCISE

Scenario: There was a campus-wide database promotion in year 2011. The promotion was discontinued in year 2012.

Sum of yearlyusers	2009	2010	2011	2012	Grand Total
ABC-CLIO eBook Collection	1154	1138	1193	1282	4767
ABI/INFORM Complete	1759	1936	1750	1945	7390
Academic OneFile	1459	1565	1505	1153	5682
Academic Search Premier	619	677	835	1016	3147
Academic Video Online	1324	1188	1322	1139	4973
Access World News	1524	1700	2030	2272	7526
AccessEmergency Medicine	1675	1983	2182	2391	8231
Accessible Archives	558	651	741	709	2659
AccessMedicine	1259	1383	1753	2002	6397
AccessPharmacy	660	625	737	794	2816
AccessScience (Encyclopedia of Science & Technology	802	924	998	1013	3737
AccessSurgery	1462	1438	1124	1078	5102
AccessUN	1227	1133	1413	1261	5034
ACLS Humanities E-Book	0	4	4	4	12
ACM Digital Library	878	947	1037	979	3841
Ad$ Summary	1112	1134	1029	1144	4419
Advertising Redbooks	233	234	229	217	913
African American Biographical Database	453	456	559	538	2006
African American Experience	1394	1382	1606	1673	6055
African American Music Reference	908	935	1037	1079	3959
African American Newspapers, 1827-1998	1313	1259	1341	1232	5145
African American Newspapers: The 19th Century	788	784	979	943	3494
African American Periodicals 1825-1995	831	976	1080	1295	4182
African Development Indicators	836	718	925	819	3298
African Writers Series	549	546	641	550	2286
African-American Poetry (1750-1900)	241	211	203	203	858
AgeLine	1411	1527	1650	1644	6232
Agricola [via EbscoHost]	1013	875	1097	982	3967
Agricola [via U.S Department of Agriculture]	1376	1435	1714	1709	6234
AGRIS	1056	1029	1109	956	4150
Alternative Press Index	1651	1717	1757	1722	6847
Alternative Press Index Archive	1073	1049	1170	1161	4453

EXERCISE

Type of Test	Use
Independent T-test	Compare the means from exactly two groups, such as the control group vs. the experimental group.
Dependent T-test	Used for before vs. after type experiments, where the same individuals are measured.
Chi-Square Test	Compare observed data with data we expect to obtain according to a specific hypothesis.
ANOVA	Compare differences between two or more groups.

Which test measure should I use to find out if there's a difference in database usage between year 2011 and 2012?

EXERCISE

P value>0.05, there's no significant difference between the two sets of data.	P value<0.05, there's a significant difference between the two sets of data.	P value>0.05, there's no significant difference between the two sets of data.

RE-EXAMINE EXAMPLES OF RESEARCH PAPERS

Citation patterns of online and print journals in the digital age

Sandra L. De Groote, MLIS, AHIP

See end of article for author's affiliation.

DOI: 10.3163/1536-5050.96.4.012

increased in 2000 and 2001 compared to 1996, although the numbers slightly decreased in 2004 and 2005. Interestingly, journals available only in print were cited on average more than the journals available in print and online formats.

A repeated measures analysis of variance (ANOVA) test was performed to examine the effects of online journals on the citation patterns of urban authors. The analysis did not find a statistically significant interaction between year of citation and print status. In other words, the findings suggested that over time, journals in print were not less likely to be cited and journals available online were not more likely to be cited, when compared to journals cited prior to the introduction of online journals. An additional ANOVA test was run on a subset (journals in the MEDLINE Abridged Index Medicus [AIM] subset) of the above journals with similar results (Table 2).

Regional campus. A total of 760 journals were cited by authors at this campus, of which 564 journals cited were excluded from the statistical analysis because

journals remained the most cited during all the years studied. There was a statistically significant interaction between the year of citation and the print status ($F=5.256$, $P<0.001$). In general, the number of cited references decreased during the study years for the print-only journals, while the cited references of the online-and-print journals increased. Cited references for the online-only journals also increased over the study period, with the exception of a decrease in 2005. Of note was the decrease overall in the number of articles cited in 2005. As also noted at the urban setting, journals available in print were cited on average more than the other journals.

An ANOVA test of the AIM journals also showed an increase in citing the online-and-print journals and a decrease in citing the print-only journals ($F=2.194$, $P<0.038$). These findings suggested that online journals have had an impact on the citation patterns of authors on the regional campus, where only a small print journal collection was available. Researchers were citing the journals available online more and citing the journals available only in print less.

De Groote, (2008) p.365

Clinical and academic use of electronic and print books: the Health Sciences Library System e-book study at the University of Pittsburgh

Barbara L. Folb, MM, MLS, MPH; Charles B. Wessel, MLS; Leslie J. Czechowski, MA, MLS

See end of article for authors' affiliations.

DOI: 10.3163/1536-5050.99.3.009

Table 3
Reported e-book use by role at University of Pittsburgh Medical Center (UPMC) or University of Pittsburgh

Affiliation and role*	Reported use of e-books n	(%)
UPMC (n=435)†		
Intern, resident, or fellow (n=91)	73	(80.2%)
Attending physician (n=71)	52	(73.2%)
Researcher (n=74)	42	(56.8%)
Other (n=25)	14	(56.0%)
Other patient care (n=35)	19	(54.3%)
Support staff (n=43)	18	(41.9%)
Nurse (n=68)	28	(41.2%)
Administrator (n=28)	8	(28.6%)
University of Pittsburgh (n=648)‡		
Postdoctoral or fellow (n=74)	54	(73.0%)
Faculty or teaching role (n=215)	139	(64.7%)
Graduate or medical student (n=205)	127	(62.0%)
Staff (n=84)	43	(51.2%)
Undergraduate (n=45)	22	(48.9%)
Other (n=25)	12	(48.0%)

* Respondents can appear in more than one category. Respondents with UPMC email addresses indicating roles at the university are included in both categories.
† χ^2=48.051, *df*=7, *P*=0.000.
‡ χ^2=13.705, *df*=5, *P*=0.018

Barriers and facilitators to e-book use

Respondent awareness and use of the e-book collection. Most respondents (n=599/914, 65.5%) recalled seeing information about e-books on the HSLS website, although slightly fewer (n=505/911, 55.4%) reported using an HSLS e-book. Use of e-books to look up brief factual information was reported by 56.6% (n=516/911), while use for in-depth study was reported by 41.9% (n=383/913).

Use and rating of e-book search tools. The utility of the 5 HSLS e-book search tools, Google Books, and the Amazon Search Within the Book feature was rated by 863 respondents, as summarized in Figure 1. The federated full-text search tool was used by the largest percent of respondents (n=580/863, 67.2%) and was rated moderately to extremely useful by 74.3% (n=431/580) who used it. Google Books was also rated as moderately to extremely useful by 74.3% (n=373/502) who used it. They gave the lowest ratings to the library catalog (PITTCat), with 61.2% (n=306/500) rating it moderately to extremely useful.

Folb, (2011) p.222

The Librarian Leading the Machine: A Reassessment of Library Instruction Methods

Katie Greer, Amanda Nichols Hess, and Elizabeth W. Kraemer

TABLE 2
Descriptive Statistics

Type	N	Mean	Std. Deviation	Std. Error Mean
Experimental Group	129	20.12	3.798	.334
Control Group	128	20.20	3.775	.334

Results

Before analysis commenced, data from students who had indicated they were underage or did not wish to have their data included in the study were removed. The remaining, anonymized data resulted in a near-equal sample size for each instructional method: 128 from the blended instruction control group and 129 from the online-only group. The authors used SPSS software to analyze the data. Table 2 provides the descriptive statistics and table 3 provides the results of an independent samples t-test.

In regard to the research question "Does the delivery format of information literacy instruction in face-to-face course sections affect attainment of student learning outcomes?" the data suggest, in this case, that it does not. The means of the two groups are nearly identical, as shown in table 2. An independent t-test, which is a generally accepted statistical measure of difference, confirmed that the groups did not differ significantly ($p = 0.88$; $p > 0.05$ confirms that there is no significant difference).

TABLE 3
Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal Variances Assumed	.835	.362	–.151	255	.880	–.07	.472	–1.002	.859
Equal Variances Not Assumed			–.151	254.999	.880	–.07	.472	–1.002	.859

Greer, (2009) p.293-294

Trends in health sciences library and information science research: an analysis of research publications in the *Bulletin of the Medical Library Association* and *Journal of the Medical Library Association* from 1991 to 2007*

Sally A. Gore, MS, MS LIS; Judith M. Nordberg, MLIS; Lisa A. Palmer, MSLS, AHIP; Mary E. Piorun, MSLS, MBA, AHIP

See end of article for authors' affiliations.

DOI: 10.3163/1536-5050.97.3.009

None	344	(72.6)	
Government	66	(13.9)	
Association	30	(6.3)	
Own institution	21	(4.4)	
Other	13	(2.7)	
Total number of authors†			=0.0019*
1	180	(38.0)	
2	142	(30.0)	
3	83	(17.5)	
4	33	(7.0)	
5	18	(3.8)	
6+	18	(3.8)	
Total number of pages‡			=0.0055*
1–4	89	(18.8)	
5–9	310	(65.4)	
10–14	68	(14.3)	
15–19	4	(0.8)	
20+	3	(0.6)	
Total number of citations§			<0.0001*
1–4	25	(5.3)	
5–9	87	(18.4)	
10–14	104	(21.9)	
15–19	91	(19.2)	
20+	167	(35.2)	

* Probability value for comparison of these results with those reported by Dimitroff [6]. $P^1 \leq 0.05$ is statistically significant.
† Mean: 2.2, median: 2, SD: 1.3.
‡ Mean: 7.0, median: 7, SD: 3.2.
§ Mean: 19.1, median: 16, SD: 15.0.

Gore, (2009) p.206

REFERENCES

1. De Groote, S. L. (2008). Citation patterns of online and print journals in the digital age. *Journal of the Medical Library Association : JMLA*, *96*(4), 362–369. http://doi.org/10.3163/1536-5050.96.4.012
2. Folb, B. L., Wessel, C. B., & Czechowski, L. J. (2011). Clinical and academic use of electronic and print books: the Health Sciences Library System e-book study at the University of Pittsburgh. *Journal of the Medical Library Association : JMLA*, *99*(3), 218–228. http://doi.org/10.3163/1536-5050.99.3.009
3. Greer, K., Hess, A. N., & Kraemer, E. W. (2015). The Librarian Leading the Machine: A Reassessment of Library Instruction Methods. *College & Research Libraries*, crl15-719.
4. Gore, Sally A, MS,M.S.L.I.S., Nordberg, J. M., M.L.I.S., Palmer, Lisa A,M.S.L.S., A.H.I.P., & Piorun, Mary E, MSLS,M.B.A., A.H.I.P. (2009). Trends in health sciences library and information science research: An analysis of research publications in the bulletin of the medical library association and journal of the medical library association from 1991 to 2007*. *Journal of the Medical Library Association, 97*(3), 203-11. Retrieved from http://libproxy.usc.edu/login?url=http://search.proquest.com/docview/203478724?accountid=14749
5. Dodge, Y. (2006) *The Oxford Dictionary of Statistical Terms*, Oxford University Press.
6. Liu, J. (2015). User assessment of search task difficulty: Relationships between reasons and ratings. *Library & Information Science Research*, *37*(4), 329-337. http://dx.doi.org/10.1016/j.lisr.2015.05.003
7. Luo, L (2016). Quantitative Analysis. Research Day at Loyola Marymount University.

RESOURCES

- Khan Academy https://www.khanacademy.org/#statistics
- Coursera Basic Statistics https://www.coursera.org/learn/basic-statistics
- Coursera Inferential Statistics https://www.coursera.org/learn/inferential-statistics-intro
- Lynda.com SPSS Statistics Essential Training https://www.lynda.com/SPSS-tutorials/SPSS-Statistics-Essential-Training/182376-2.html?org=usc.edu