Statistics: A Vital Tool in Research

R. Jeanette O'Hara Hines

Dept. of Statistics and Actuarial Sci.

Director of the SCS

Steps involved in scientific research:

PPDAC

is a systematic framework for solving problems and for acquiring knowledge

Problem

Plan

Data

Analysis

Conclusion

Problem: a clear statement of what you are trying to learn, of what questions you want to answer

- **Are 'clickers' an effective teaching tool?**
- **Do 'case studies' help students learn concepts better?**
- **Is a proposed 'blended' method of teaching effective?**

Understanding the Problem

Need to formulate the problem well

What is your research ‘hypothesis’?

What conclusions are you trying to draw?

About what group of things or people do you want the conclusions to apply?

What will you measure?

What is the question you are trying to answer?

Some of the above questions will be answered in detail in the Plan.

Plan: the procedures you will use to carry out the study; how you measure and collect data

There is a lot of work involved at this step, and you may have to go back to the Problem statement to clarify or to revise it.

Statistical planning is already needed at the 'Planning' stage:

- **how will the info needed be obtained?**
- **how will you measure to get the info?**
- **how will the data be collected?**
- **what and who will be measured?**
- **what statistical tools are needed to examine data?**
- **will data gathered and conclusions reached answer questions?**

Is study ‘observational’?

- an ‘observational’ study is one in which the researcher just observes and records the info. Association only; cannot establish ‘causation’. Surveys are often observational studies

Or will an <u>experiment</u> be used?

- a simple experiment is a study in which the researcher deliberately manipulates variables (like teaching style/method, learning environment), randomly assigns objects/subjects to treatments and then records the resulting info. Can establish 'causation'.

Unit: basic element on which measurements are made; sampling or experimental unit

Target pop^{n} : well-defined set of units about which you want to draw conclusions

Study pop^{n} : well-defined set of units about which you can draw conclusions

Sample: subset or part of study population selected in some manner

Variable: quantity or characteristic measured on each unit

Types of variables

Categorical

Yes/No; Male/Female

Learning style; Area of study

Ordered Categorical

Age category; Education level; Strongly agree to strongly disagree; Income level

Numeric

Age; Weight; Income; Test score

Pop^{n} attribute: a characteristic of a variable in the study pop^{n}; for example, the percent of yeses, the average test score, the spread of test scores

For a survey:

What group ('target' population) do you want to survey? What group ('study' population) can you actually survey? How will you get a 'sample' from this latter group?

What questions do you want to ask? Are they clear and will they measure the info you want?

Pre-test your questionnaire, and do it again.

Think about the responses you might get and whether they will address your questions in the Problem statement.

Data: the data are collected according to the Plan

Analysis: the data are summarized and analyzed to answer the questions posed

Examples of some data collected, summarized and analyzed

1-way Tables

Gender

Male	Female	Total
102 (35.5%)	185 (64.5%)	287

Score on a 1-7 ordinal scale (say, strongly agree to strongly disagree)

1	2	3	4	5	6	7	Total
1	0	1	58	141	80	6	287

2-way Tables

	Yes	No	Total
Male	48	54	102
Female	75	110	185
Total	123	164	287

Research Question:

Is there an association between Gender and Yes/No question? Do Females have a higher or lower chance of saying Yes compared to Males?

Statistical Statements or Hypotheses:

Null Hypothesis: There is no association between Gender and Yes/No question

Alternative Hypothesis: There is an association between Gender and Yes/No question

Test: Chi-square test

Results: $\chi^2 = 1.14$ with 1 df; P = 0.29

Note: P = P-value = measure (0 to 1) of agreement of data with null hypothesis

P large (say > 0.10) ➜ no evidence that data disagree with null hypothesis

P small (say < 0.05) ➜ evidence that data disagree with null hypothesis; alternative more reasonable

Statistical conclusion: No evidence of an association between Gender and Yes/No question

OR

No evidence that Females have a higher or lower chance of saying Yes compared to Males

	Yes	No	Total
Male	47.1%	52.9%	102
Female	40.5%	59.5%	185
Total	123	164	287

Another example

Null Hypothesis: There is no association between Gender and 1-7 ordinal scale question

Alternative Hypothesis: There is an association between Gender and 1-7 ordinal scale question OR Males have a different distribution of ‘1-7' scores than Females

	1	2	3	4	5	6	7	Total
M	1	0	1	21	47	31	1	102
F	0	0	0	37	94	49	5	185
T	1	0	1	58	141	80	6	287

	1-4	5	6-7	Total
M	23	47	32	102
F	37	94	54	185
Total	60	141	86	287

Test: Chi-square test

Results: $\chi^2 = 0.61$ with 2 df; P = 0.74

Statistical conclusion: No evidence of an association between Gender and '1-7' ordinal scale question

OR

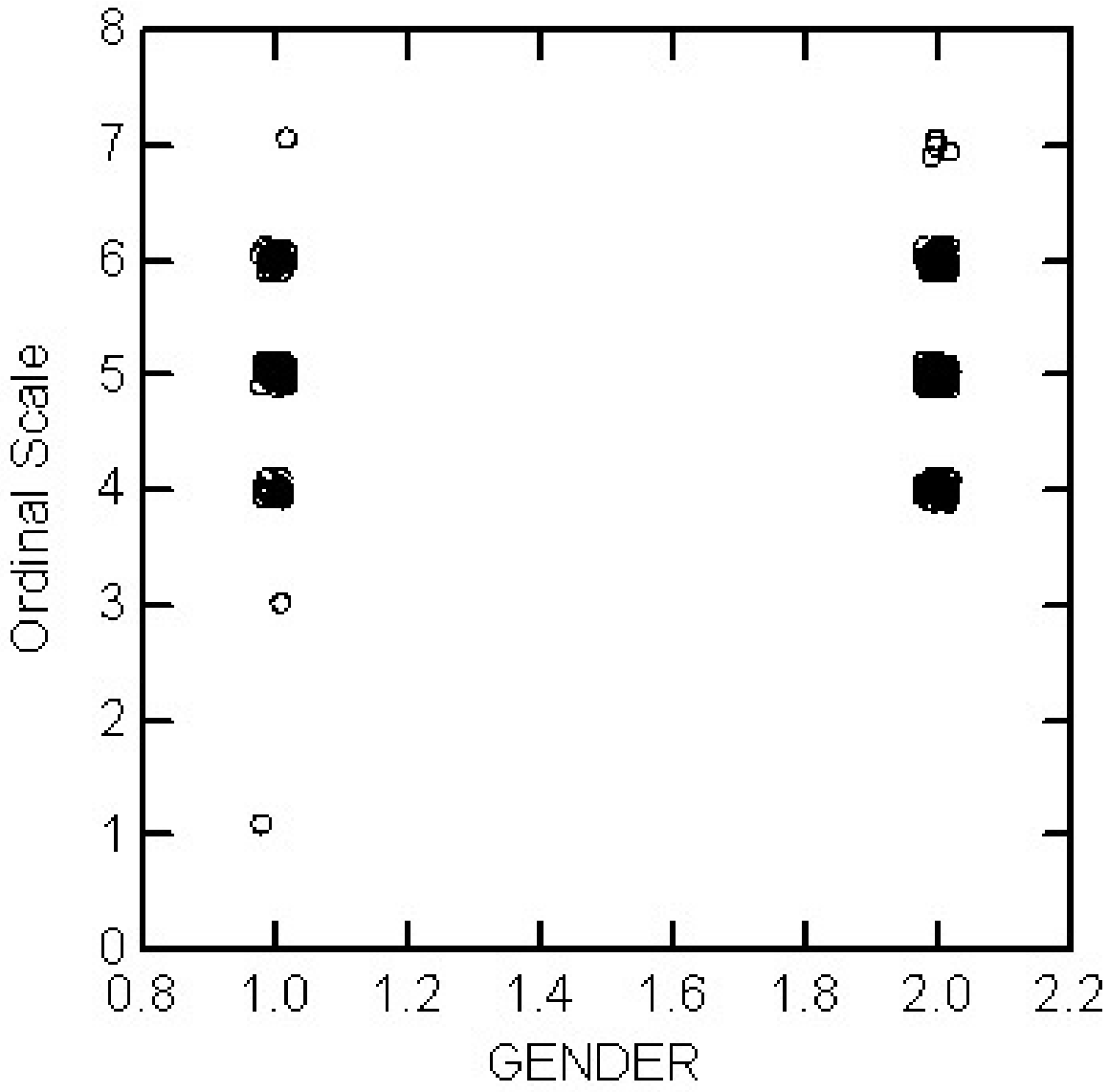
No evidence that Females and Males have a different distribution of '1-7' scores

Cell Percentages

	1-4	5	6-7	Total
M	22.5	46.1	31.4	102
F	20.0	50.8	29.2	185
Total	60	141	86	287

Another way to analyze data, using 1-7

A non-parametric t-test (Mann-Whitney) tests for a difference in MEDIANS between M and F. Scores are replaced by ranks (1 to 287) with ties adjusted for.


Ordinal Scale
0
1
2
3
4
5
6
7
8
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
GENDER

Median Score:

Males: 5.0 **Females: 5.0**

Null hypothesis: the median score for the females is the same as that for Males

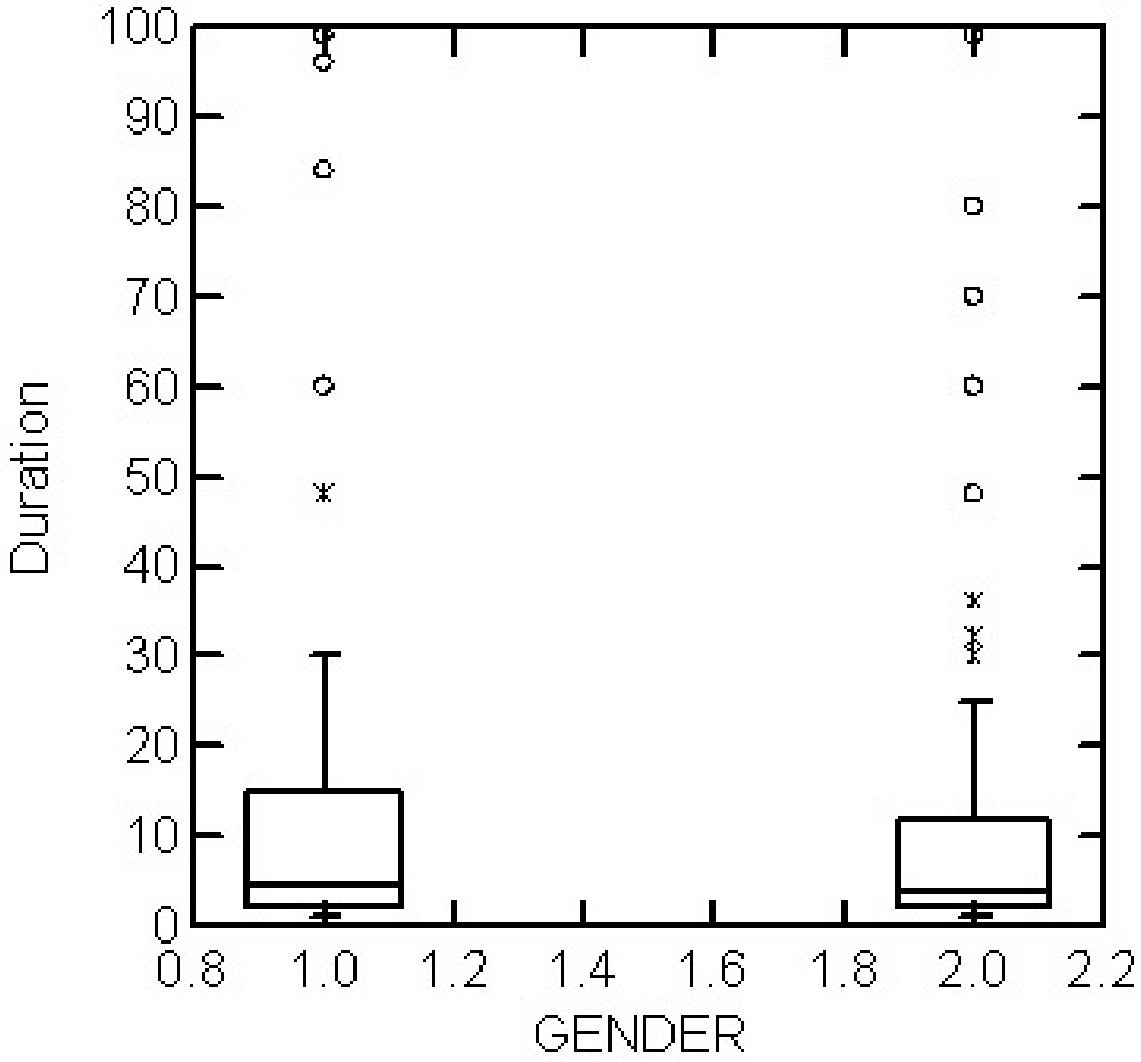
Results: P = 0.88 ➜ no evidence of a difference in median score between males and females

Another example

Duration (of something) measured in months ranges from 1 to 99 (?).

You are interested in testing if ‘average’ duration differs between Females and Males

Null hypothesis: the ‘average’ duration for the Females is the same as that for Males

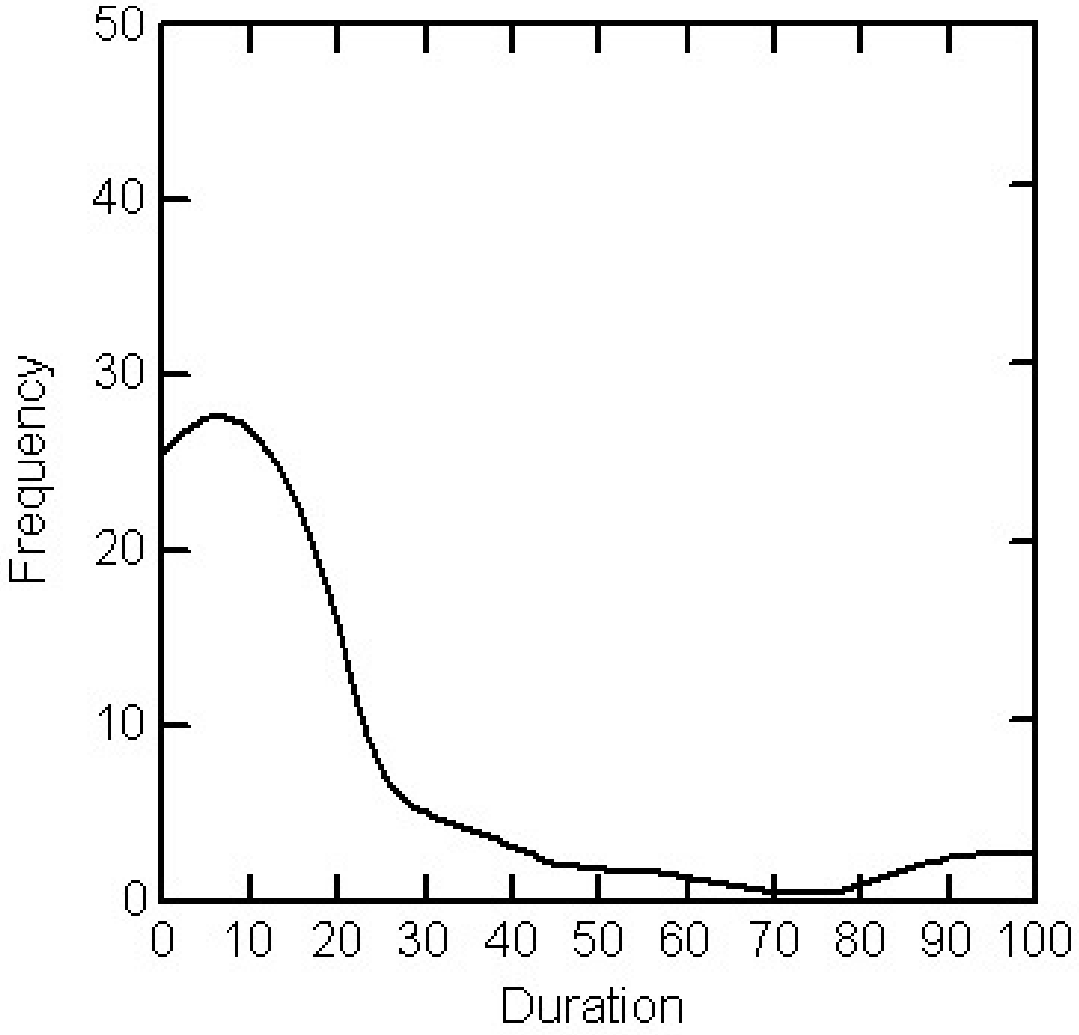

100
90
80
70
60
50
40
30
20
10
0
Duration
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
GENDER

Different possible tests - which one is best?

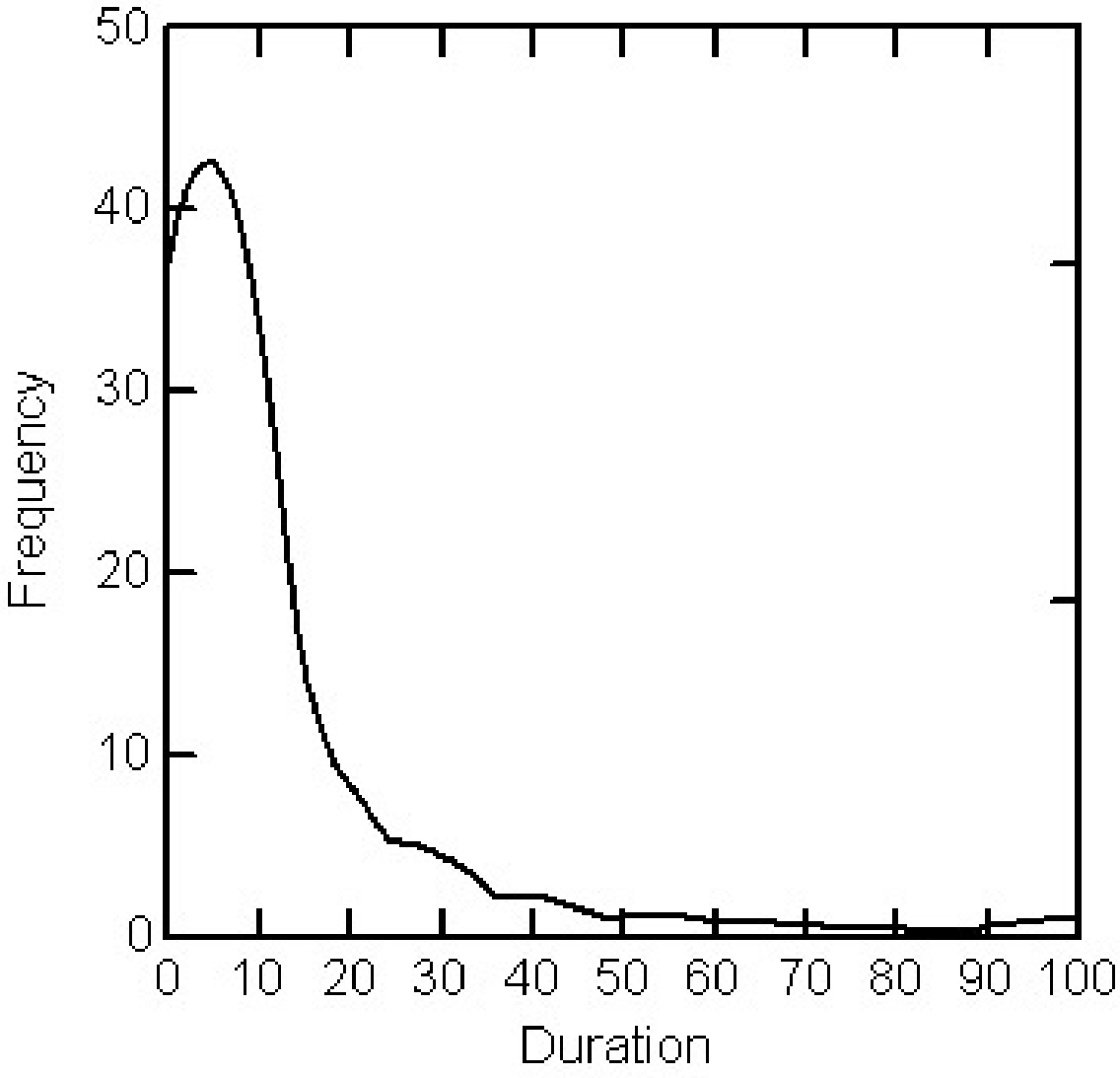
1. parametric t-test (assumes approximate normality of variable [duration] in each group)

2. non-parametric t-test (loss of power if 1 more appropriate) and tests for difference in median

Gender 1, n = 102



Gender 2, n = 185



Variable = Duration (in months)

Gender	1	2
Mean	16.2	10.6
Median	4.5	4.0
SD	26.5	17.3
n	102	185

2-sample t-tests

parametric t-test

pooled variance: P = 0.03

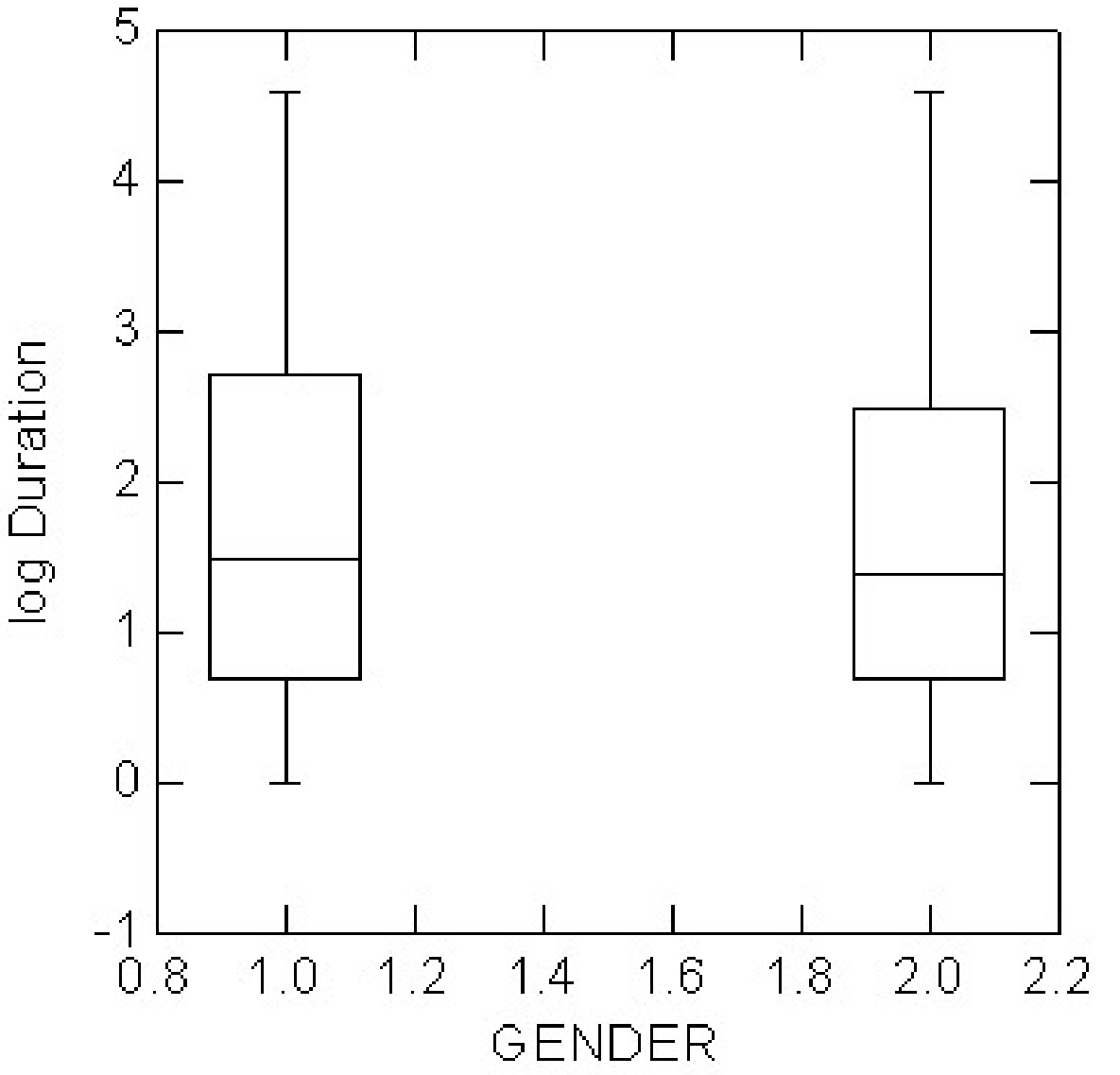
separate variance: P = 0.06

non-parametric t-test (Mann-Whitney)

P = 0.20

Which one do you use and what is your conclusion?

Could also log duration and run parametric t-test.



Variable = log Duration

Gender	1	2
Mean	1.81	1.34
SD	1.34	1.19
n	102	185

t-test

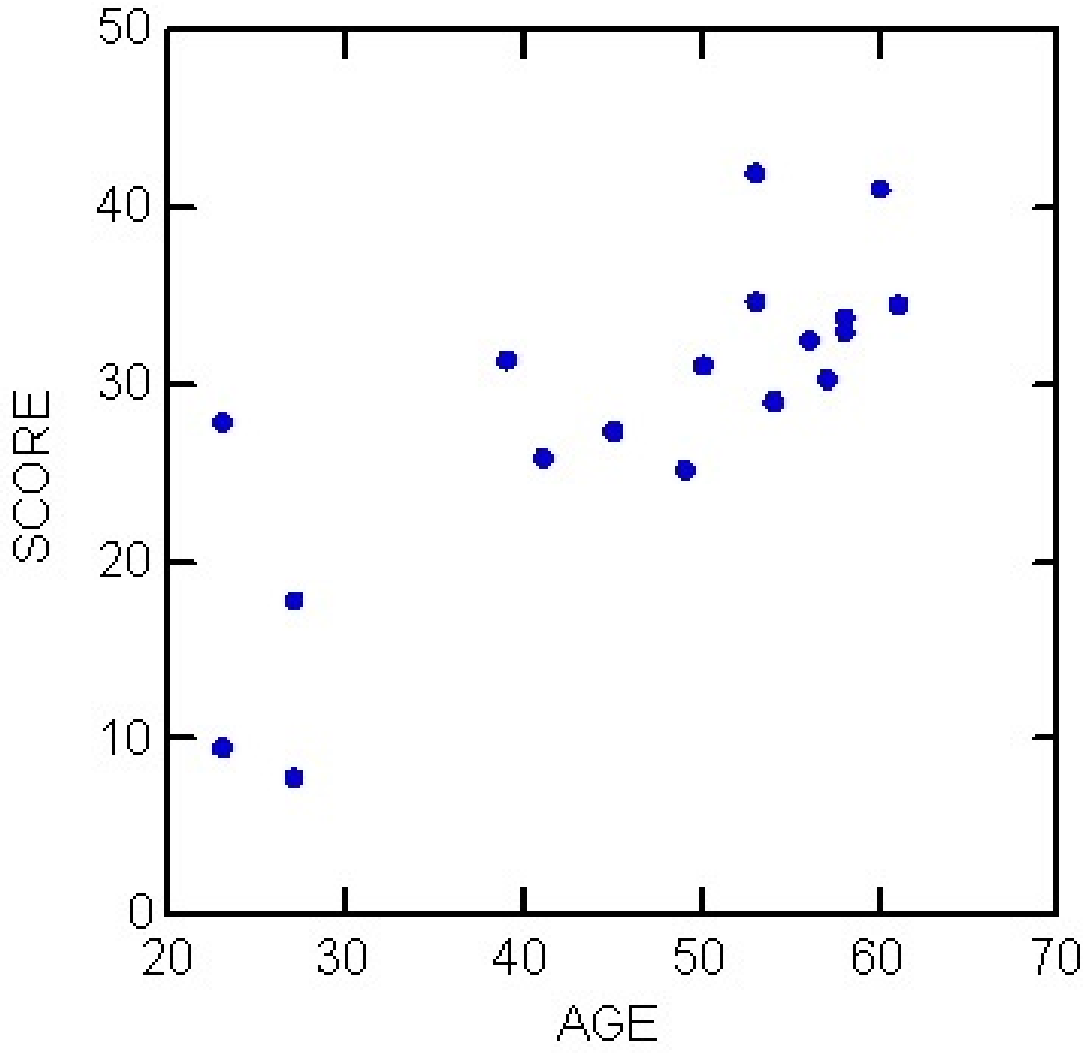
pooled variance: P = 0.12

separate variance: P = 0.13

Interested in testing if two variables are related ➜ correlation (perhaps simple linear regression if prediction equation needed)

Is score linearly related to age? n = 18

Null hypothesis: score and age are NOT linearly related


50
40
30
20
10
0
SCORE
20
30
40
50
60
70
AGE

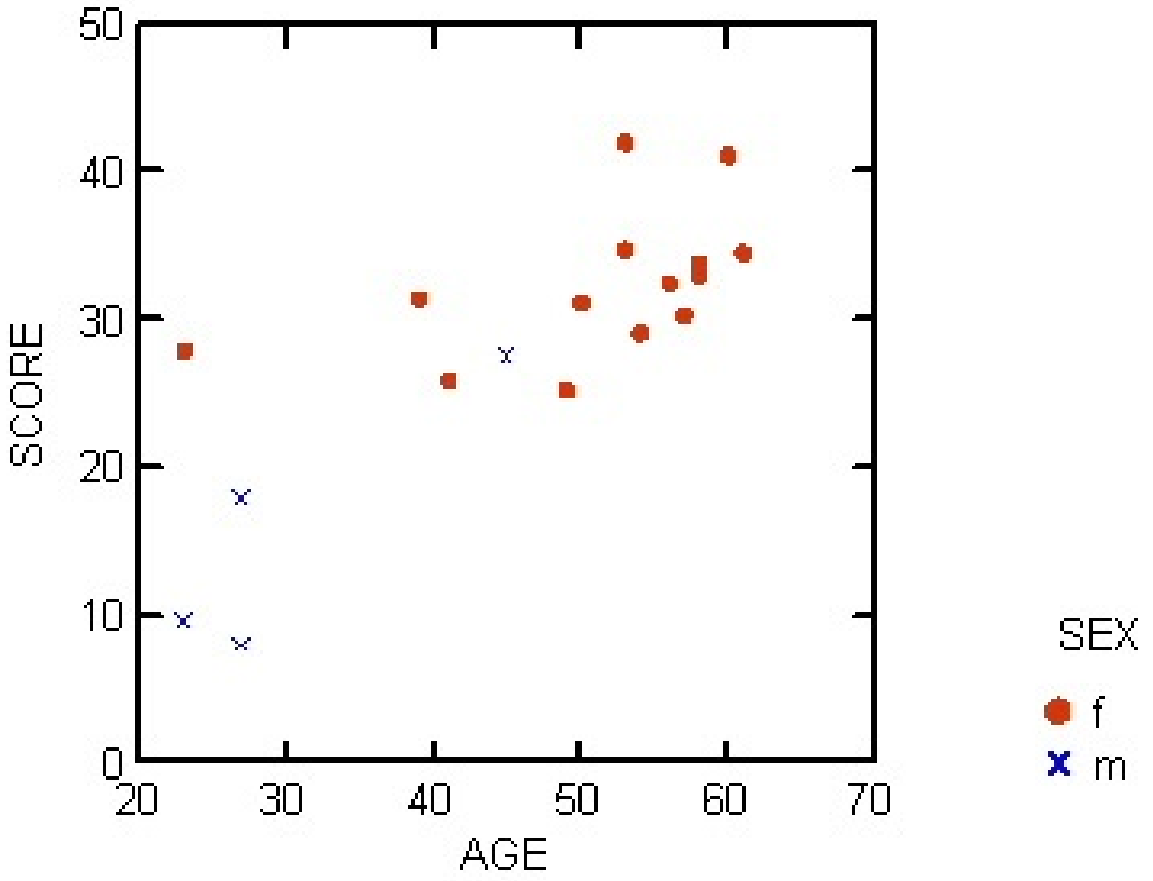
correlation r = 0.79 with 16 df; P = 0.000

Conclusion: there is a strongly significant correlation between score and age

Note: strongly significant does NOT mean strongly correlated

How significant is determined by n; how strongly correlated is determined by r^2

Here $r^2 = 0.79^2 = 0.62$ OR 62%


50
40
30
20
10
0
SCORE
20
30
40
50
60
70
AGE
SEX
f
m

Just Females n = 14

r = 0.51 with 12 df P = 0.07

r^2 = 0.26 OR 26%

Conclusions: Conclusions are drawn about what has been learned

You summarize the statistical results in the study terminology

Other commonly used statistical tools:

comparing the averages of more than 2 groups - ANOVA

relating a single response variable to two or more covariates/explanatory variables - multiple regression

For more statistical help, please come and see one of us at the Statistical Consulting Service

http://www.stats.uwaterloo.ca/stats_navigation/Consulting/StatConsulting.shtml

OR search 'Statistical Consulting' on UW website

Example: Blended Learning Survey

Problem: to measure the pre- and post-attitudes of students to the 'blended' mode of course delivery for future use of such courses

blended approach to learning involves integrated face-to-face and online environments

Plan: Prepare surveys and pre-test them. Ask students in Pharm 220 to complete pre- and post- surveys on-line. Surveys measure attitudes about 'blended' mode of course delivery. Students are encouraged to fill out surveys but are not obligated to. E-mails and in class announcements and reminders to fill out survey are given. Students told that study has approval from the Office of Research Ethics.

Variables measured on surveys: 5-point scale questions measuring pre- and post- levels of enthusiasm for blended course, yes/no pre- questions about having taken a blended and fully on-line course before, 5-point scale post questions measuring level of learning/understanding given by various aspects of course and level of satisfaction, some open-ended questions

The Planning stage is also where the researchers should outline what statistical tools they will use to analyze the incoming data, and how statistical conclusions will translate into practical ones.

Researchers must make sure that data being gathered will address their intended questions and the data is actually analyzable.

Data: data are gathered from pre- and post- surveys. Pre- and post-surveys for each student are matched by having the student create an identifier which he/she writes on both surveys.

Analysis:

Statistical tools

to test for pre- to post- changes in 5-point scales - probably a non-parametric paired comparison test such as a signed-rank test

to summarize yeses/nos for pre- questions on blended and fully on-line courses and 5-point post scale questions measuring level of learning/understanding, enthusiasm - tables or bar charts

Conclusions:

Translate statistical conclusions into research conclusions using the language of the survey

Unit: an undergraduate university student

Target pop^{n}: all undergraduate university students

Study pop^{n}: university students taking Pharm 220 ➜ sample consists of volunteers since students can choose to answer surveys or not

Blended Learning Survey

is sort of an experimental study

in which pre and post surveys are used to describe/summarize students' changes (pre to post course) in attitudes to the 'blended' form of course delivery; changes will be attributed to course delivery but could possibly be due to some other change over time that could not be controlled