



# Statistical Tests

## Which test should I Use?

Miriam Nuño, PhD

Associate Professor in Biostatistics and In Residence of Surgery  
Department of Public Health Sciences  
UC Davis School of Medicine

September 11, 18, 2019

- UC Davis Health Clinical and Translational Science Center
- UC Davis Health Mind Institute
- UC Davis Health Comprehensive Cancer Center
- UC Davis Environment Health Sciences Center

**We are video recording this seminar  
so please hold questions until the end.**

**Thanks**



# Seminar Objectives

- Learn how to choose the appropriate test for your data
- Provide an overview of different types of tests
- Learn how to perform the tests on SAS
- Next month's seminar will cover scenarios when your data don't meet the assumptions of the parametric test



# Types of Data

- **Nominal data:** Observations fall into categories that can't be ordered.  
(e.g. Mediterranean diet, Atkins diet, South Beach diet)
- **Ordinal data:** Observations fall into ordered categories.  
(e.g. underweight, normal weight, overweight, obese, morbidly obese)
- **Interval scale data:** Observations are ordered, distance between possible values is meaningful, but no true “zero” point (impossible to compute ratios)  
(e.g. Temperature: there is a zero but it has meaning, 20° is not twice as hot as 10°)
- **Ratio scale data:** Observations are ordered, distance is meaningful, and a floor of “true zero” (complete absence of anything, no negative numbers) makes ratios meaningful (e.g. weight, height, age)

# Determine test to use

- **What type of variable is the outcome?**
  - Continuous/Numeric
    - e.g. height (inches), weight (pounds)
  - Categorical
    - e.g. Gender, Race
  - Survival, time until an event occurs
    - e.g. Time until tumor recurrence, Time until cardiovascular death after some treatment intervention
  
- **What type of variable is the predictor?**
  - Categorical, continuous

# Determine test to use

	PREDICTORS			
	Categorical			Continuous
OUTCOME	1 group	2 groups	>2 groups	
Continuous	One-sample t-test	Two-sample t-test (dependent, independent)	ANOVA	Linear regression
Categorical (binary)	One-sample proportion test	Two-sample proportion test (dependent, independent)	Chi-square test	Logistic Regression
Survival	Kaplan-Meier Estimate	Log Rank Test	Log Rank Test	Proportional hazards regression

# Tests and Examples

TEST	EXAMPLE
One-sample t-test	Joint position sense
Two-sample t-test	A small clinical study: weight
Paired t-test	Study: bran in the treatment of diverticulosis
One-sample proportion test	Test the proportion of babies
Two-sample proportion test	Gender differences relative to smoking behavior
McNemar's test	Treatments for athlete's foot
Kaplan-Meier Estimate	HMO-HIV+ Study
Log rank test	HMO-HIV+ Study

# Continuous outcome



# Tests and Examples

Test	Example
<b>One-sample t-test</b>	<b>Joint position sense</b>
Two-sample t-test	A small clinical study: weight
Paired t-test	Study: bran in the treatment of diverticulosis
One-sample proportion test	Test the proportion of babies
Two-sample proportion test	Gender differences relative to smoking behavior
McNemar's test	Treatments for athlete's foot
Kaplan-Meier Estimate	HMO-HIV+ Study
Log rank test	HMO-HIV+ Study

# One-sample t-test

## Example: joint position sense

- Investigate ability to know what position our joints are without looking or touching
- Test whether people over- or underestimate their knee angle
- Subjects bend the knee to a  $120^\circ$  angle for a few seconds, then return the knee to a  $90^\circ$  angle. Then each person bend their knee to the  $120^\circ$  angle again
- The measurement variable is the angle of the knee, and the theoretical expectation from the null hypothesis is 120

# One-sample t-test

Example: joint position sense

- **One group**
  - 16 subjects
- **Outcome of interest**
  - the angle of the knee
- **Continuous outcome with one group**
  - One-sample t-test

# One-sample t-test

## Hypothesis being tested

- **The null hypothesis**

- people don't over- or underestimate their knee angle

$$H_0: \mu = 120$$

- **The alternative hypothesis**

- people over- or underestimate their knee angle

$$H_1: \mu \neq 120$$

# One-sample t-test

## Assumptions

- The data are continuous
  - The data following a normal distribution
  - Samples are independent and random
  - The population standard deviation is unknown
- 

# One-sample t-test

## SAS code

```
data onesamletttest;
    input angle @@;
    datalines;
120.6 116.4 117.2 118.1 114.1 112.1 115.7 112.9
116.9 113.3 121.1 116.9 117.0 114.0 123.0 119.1
;
run;

title 'One Sample T-test';
proc ttest data=onesamletttest h0=120 plots(showh0) sides=2 alpha=0.05;
    var angle;
run;
```

# One-sample t-test

## SAS output

### One Sample T-test

The TTEST Procedure

Variable: angle

N	Mean	Std Dev	Std Err	Minimum	Maximum
16	116.8	3.1049	0.7762	112.1	123.0

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
116.8	115.1	118.4	3.1049	2.2936	4.8055

DF	t Value	Pr >  t
15	-4.15	0.0008

# List of tests and examples

TEST	EXAMPLE
One-sample t-test	Joint position sense
<b>Two-sample t-test</b>	<b>A small clinical study: weight</b>
Paired t-test	Study: bran in the treatment of diverticulosis
One-sample proportion test	Test the proportion of babies
Two-sample proportion test	Gender differences relative to smoking behavior
McNemar's test	Treatments for athlete's foot
Kaplan-Meier Estimate	HMO-HIV+ Study
Log rank test	HMO-HIV+ Study

# Two-sample t-test

Example: A small clinical study: weight

- **Collect weight information on patients**
- **Two groups (independent)**
  - 1) Male and 2) Female
- **Outcome of interest**
  - difference in weight between men and women
- **Continuous outcome with two groups**
  - Two-sample t-test

# Two-sample t-test

## Hypothesis being tested

- **The null hypothesis**

- The mean weight of male study patients is not different from that of the female study patients

$$H_0: \mu_1 = \mu_2$$

- **The alternative hypothesis**

- The mean weight of male study patients is different from that of the female study patients

$$H_1: \mu_1 \neq \mu_2$$

# Two-sample t-test

## Assumptions

- The data are continuous
- The data in each group following a normal distribution
- The two samples are independent
- Both samples are simple random samples from their respective populations

# Two-sample t-test

## SAS code

```
data twosamplesttest;
  input sex $ weight @@;
  datalines;
F 85.0 F 105.0 F 108.0 F 92.0 F 112.5
F 112.0 F 104.0 F 94.5
M 112.0 M 114.0 M 140.0 M 107.5 M 87.0
;
title 'Two Sample T-Test';
proc ttest data=twosamplesttest sides=2 alpha=0.05;
  class sex; /* defines the grouping variable */
  var weight; /* variable whose means will be compared */
run;
```

# Two-sample t-test

## SAS output

### Two Sample T-Test

The TTEST Procedure  
Variable: weight

sex	N	Mean	Std Dev	Std Err	Minimum	Maximum
F	8	101.6	10.0241	3.5440	85.0000	112.5
M	5	112.1	18.9288	8.4652	87.0000	140.0
Diff (1-2)		-10.4750	13.9368	7.9452		

sex	Method	Mean	95% CL Mean	Std Dev	95% CL Std Dev
F		101.6	93.2447 110.0	10.0241	6.6277 20.4017
M		112.1	88.5968 135.6	18.9288	11.3409 54.3930
Diff (1-2)	Pooled	-10.4750	-27.9623 7.0123	13.9368	9.8728 23.6630
Diff (1-2)	Satterthwaite	-10.4750	-33.5149 12.5649		

Method	Variances	DF	t Value	Pr >  t
Pooled	Equal	11	-1.32	0.2142
Satterthwaite	Unequal	5.4298	-1.14	0.3015

Test statistics

Test for equal variance

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	4	7	3.57	0.1371

P-value for the test assumed unequal variances.

# Tests and Examples

Test	Example
One-sample t-test	Joint position sense
Two-sample t-test	A small clinical study: weight
<b>Paired t-test</b>	<b>Study: bran in the treatment of diverticulosis</b>
One-sample proportion test	Test the proportion of babies
Two-sample proportion test	Gender differences relative to smoking behavior
McNemar's test	Treatments for athlete's foot
Kaplan-Meier Estimate	HMO-HIV+ Study
Log rank test	HMO-HIV+ Study

# Paired t-test

**Example: bran in the treatment of diverticulosis**

- Does transit time through the alimentary canal differ if bran is given in the same dosage in three meals during the day (treatment A) or in one meal (treatment B)?
- A random sample of patients with disease of comparable severity and aged 20-44 is chosen

# Paired t-test

**Example: bran in the treatment of diverticulosis**

- **The two treatments administered on two successive occasions**
- **Two groups (dependent)**
  - 1) Treatment A
  - 2) Treatment B
- **Outcome of interest**
  - alimentary transit times
- **Continuous outcome with two paired measurements on the same subject**
  - paired t-test

# Paired t-test

## Hypothesis being tested

- **The null hypothesis**
  - There is no difference in mean transit times on between these two treatments

$$H_0: \mu_A = \mu_B$$

- **The alternative hypothesis**
  - There is a difference in mean transit times between these two treatments

$$H_1: \mu_A \neq \mu_B$$

# Paired t-test

## Assumptions

- The data are continuous
- The data, more specifically the differences for the matched-pairs, follow a normal probability distribution
- The sample of pairs is a simple random sample from its population.

# Paired t-test

## SAS code

```
data pairedttest;
  input A B @@;
  datalines;
63 55  54 62  79 108  68 77  87 83  84 78
92 79  57 94  66 69  53 66  76 72  63 77
;
title 'Paired T-Test';
proc ttest data=pairedttest sides=2 alpha=0.05;
  paired A*B;
run;
```

# Paired t-test

## SAS output

### Paired T-Test

The TTEST Procedure  
Difference: A - B

N	Mean	Std Dev	Std Err	Minimum	Maximum
12	-6.5000	15.1448	4.3719	-37.0000	13.0000

Mean	95% CL Mean	Std Dev	95% CL Std Dev		
-6.5000	-16.1225	3.1225	15.1448	10.7285	25.7139

DF	t Value	Pr >  t
11	-1.49	0.1652

# Categorical Outcome



# Tests and Examples

TEST	EXAMPLE
One-sample t-test	Joint position sense
Two-sample t-test	A small clinical study: weight
Paired t-test	Study: bran in the treatment of diverticulosis
<b>One-sample proportion test</b>	<b>Test the proportion of babies</b>
Two-sample proportion test	Gender differences relative to smoking behavior
McNemar's test	Treatments for athlete's foot
Kaplan-Meier Estimate	HMO-HIV+ Study
Log rank test	HMO-HIV+ Study

# One-sample proportion test

Example: Test the proportion of babies

- **Sample 28 babies from a group under certain treatment**
- **One group**
- **Outcome of interest**
  - Gender of baby
- **Categorical outcome with one group**
  - One-sample proportion test

# One-sample proportion test

## Hypothesis being tested

- **The null hypothesis**

- The proportion of male babies is no different from 50%

$$H_0: p = 0.5$$

- **The alternative hypothesis**

- The proportion of male babies is different from 50%

$$H_1: p \neq 0.5$$

# One-sample proportion test

## Assumptions

- The data are a simple random sample from the population of interest
- The sample size  $n$  is large enough so that numbers of observations in each label are 10 or more.

# One-sample proportion test

## SAS code

```
data onesampleproportiontest;
  input Gender $ @@;
  datalines;
M M M M M M M M M M M M M M M
F F F F F F F F F F F F F
;
title 'One-sample proportion test';
proc freq data=onesampleproportiontest;
  tables Gender / binomial(p=0.5);
run;
```

# One-sample proportion test

## SAS output

### One-sample proportion test

The FREQ Procedure

Gender	Frequency	Percent	Cumulative Frequency	Cumulative Percent
F	13	46.43	13	46.43
M	15	53.57	28	100.00

#### Binomial Proportion

Gender = F

Proportion	0.4643
ASE	0.0942
95% Lower Conf Limit	0.2796
95% Upper Conf Limit	0.6490
Exact Conf Limits	
95% Lower Conf Limit	0.2751
95% Upper Conf Limit	0.6613

#### Test of H0: Proportion = 0.5

ASE under H0	0.0945
Z	-0.3780
One-sided Pr < Z	0.3527
Two-sided Pr >  Z	0.7055

Sample Size = 28

# Tests and Examples

TEST	EXAMPLE
One-sample t-test	Joint position sense
Two-sample t-test	A small clinical study: weight
Paired t-test	Study: bran in the treatment of diverticulosis
One-sample proportion test	Test the proportion of babies
<b>Two-sample proportion test</b>	<b>Gender differences relative to smoking behavior</b>
McNemar's test	Treatments for athlete's foot
Kaplan-Meier Estimate	HMO-HIV+ Study
Log rank test	HMO-HIV+ Study

# Two-sample proportion test

Example: Gender differences relative to smoking behavior

- According to an American Cancer Society report, more men than women smoke
- In a random sample of 200 males and 200 females, 62 of the males and 54 of the females were smokers
- Is there sufficient evidence to conclude that the proportion of male smokers different from the proportion of female smokers?



# Two-sample proportion test

Example: Gender differences relative to smoking behavior

- **Two groups**
  - 1) male 2) female
- **Outcome of interest**
  - Rate of smokers (proportion)
- **Categorical outcome with two groups**
  - Two sample proportion test (equivalent to chi-square test)
  - Use Fisher exact test for small sample



# Two-sample proportion test

## Hypothesis being tested

- **The null hypothesis**
  - the proportion of male smokers is no different from the proportion of female smokers
- **The alternative hypothesis**
  - the proportion of male smokers not equal to the proportion of female smokers

# Two-sample proportion test

## Assumptions

- The data are a simple random sample from the population of interest
- A minimum of 10 successes and 10 failures in each group
  - Use Fisher exact test for small numbers
- The two groups that are being compared must be unpaired and unrelated

# Two-sample proportion test

## SAS code

```
data twoindependentproptest;
  input Gender $ Smoker Total;
    Response="Smoker"; Count=Smoker;          output;
    Response="Nonsmoker"; Count=Total-Smoker; output;
  datalines;
Men      62 200
Women   54 200
;
title 'Two independent samples proportion test';
proc freq data=twoindependentproptest;
  weight Count;
  table Gender * Response / chisq riskdiff;
run;
```

# Two-sample proportion test

## SAS output

### Two independent samples proportion test

The FREQ Procedure

Frequency Percent Row Pct Col Pct	Table of Gender by Response			
	Gender	Response		Total
		Nonsmo	Smoker	
Men		138	62	200
		34.50	15.50	50.00
		69.00	31.00	
		48.59	53.45	
Women		146	54	200
		36.50	13.50	50.00
		73.00	27.00	
		51.41	46.55	
Total		284	116	400
		71.00	29.00	100.00

#### Statistics for Table of Gender by Response

Statistic	DF	Value	Prob
Chi-Square	1	0.7771	0.3780
Likelihood Ratio Chi-Square	1	0.7775	0.3779
Continuity Adj. Chi-Square	1	0.5949	0.4405
Mantel-Haenszel Chi-Square	1	0.7751	0.3786
Phi Coefficient		-0.0441	
Contingency Coefficient		0.0440	
Cramer's V		-0.0441	

#### Fisher's Exact Test

Cell (1,1) Frequency (F)	138
Left-sided Pr <= F	0.2203
Right-sided Pr >= F	0.8393
Table Probability (P)	0.0596
Two-sided Pr <= P	0.4406

# List of tests and examples

TEST	EXAMPLE
One-sample t-test	Joint position sense
Two-sample t-test	A small clinical study: weight
Paired t-test	Study: bran in the treatment of diverticulosis
One-sample proportion test	Test the proportion of babies
Two-sample proportion test	Gender differences relative to smoking behavior
<b>McNemar's test</b>	<b>Treatments for athlete's foot</b>
Kaplan-Meier Estimate	HMO-HIV+ Study
Log rank test	HMO-HIV+ Study

# McNemar's test

## Example: Treatments for athlete's foot

- Assume that each subjects has athlete's foot on each foot
- Each subject is given a treatment X on one foot and Y on the other foot
- Because left and right feet of the same subject are not independent, contingency test cannot be used

# McNemar's test

## Example: Treatments for athlete's foot

- Table of treatment X by treatment Y

---

		Treatment Y		
		cured	Not cured	Total
Treatment X	cured	12	8	20
	Not cured	40	20	60
	Total	52	28	80

---

# McNemar's test

## Example: Treatments for athlete's foot

- **Two groups**
  - 1) treatment X 2) treatment Y
- **Outcome of interest**
  - Foot cured or not
- **Categorical outcome with two dependent groups**
  - McNemar's test

# McNemar's test

## Hypothesis being tested

- **The null hypothesis**

- The paired sample proportions are equal and no (significant) change has occurred.

$$H_0: p_b = p_c$$

- **The alternative hypothesis**

- The paired sample proportions are not equal

$$H_0: p_b \neq p_c$$


# McNemar's test

Example: Treatments for athlete's foot

- Table of treatment X by treatment Y

---

		Treatment Y		
		cured	Not cured	Total
Treatment X	cured	a	b	a+b
	Not cured	c	d	c+d
	Total	a+c	b+d	n

---

# McNemar's test

## Assumptions

- The sample was randomly selected
- The sample data consists of matched pairs
- There are 2 variables each with two categories
- The frequencies are big enough such that  $b+c \geq 10$

# McNemar's test

## SAS code

```
proc format;
    value $result 'c'='cured'
                 'nc'='Not cured';
data twodependentproptest;
    infile "/folders/myfolders/Presentation/athelete.csv"
    dlm=',' firstobs=2;
    input subject treatX $ treatY $;
    format treatX treatY $result. ;
run;
proc freq data=twodependentproptest;
    title "McNemar's test for Paired Samples";
    tables treatX*treatY /agree expected norow nocol nopercnt;
run;
```

# McNemar's test

## SAS output

### McNemar's test for Paired Samples

The FREQ Procedure

Frequency  
Expected

Table of treatX by treatY			
treatX	treatY		Total
	cured	Not cured	
cured	12	8	20
Not cured	13	7	20
Total	40	20	60
Total	39	21	60
Total	52	28	80

Statistics for Table of treatX by treatY

McNemar's Test	
Statistic (S)	21.3333
DF	1
Pr > S	<.0001

Simple Kappa Coefficient	
Kappa	-0.0435
ASE	0.0821
95% Lower Conf Limit	-0.2044
95% Upper Conf Limit	0.1174

Sample Size = 80

# Survival Outcome



# Tests and Examples

TEST	EXAMPLE
One-sample t-test	Joint position sense
Two-sample t-test	A small clinical study: weight
Paired t-test	Study: bran in the treatment of diverticulosis
One-sample proportion test	Test the proportion of babies
Two-sample proportion test	Gender differences relative to smoking behavior
McNemar's test	Treatments for athlete's foot
<b>Kaplan-Meier Estimate</b>	<b>HMO-HIV+ Study</b>
Log rank test	HMO-HIV+ Study

# Kaplan-Meier Estimate

- Also known as the product limit estimator
- A non-parametric statistic used to estimate the survival function from lifetime data
- Often used to measure the fraction of patients living for a certain amount of time after treatment

# Kaplan-Meier Estimate

## Example: HMO-HIV+ study

- List of variables:

Variable	Description	Codes/Units
ID	Subject ID Code	1-100
TIME	Survival Time	survival time (in months)
CENSOR	Follow-Up Status	1 = Death due to AIDS or AIDS related factors 0 = Alive at study end or lost to follow-up
DRUG	History of IV Drug Use	0 = No 1 = Yes

# Kaplan-Meier Estimate

Example: HMO-HIV+ study

- First five observations

ID	time	ensor	drug
1	5	1	0
2	6	0	1
3	8	1	1
4	3	1	1
5	22	1	0

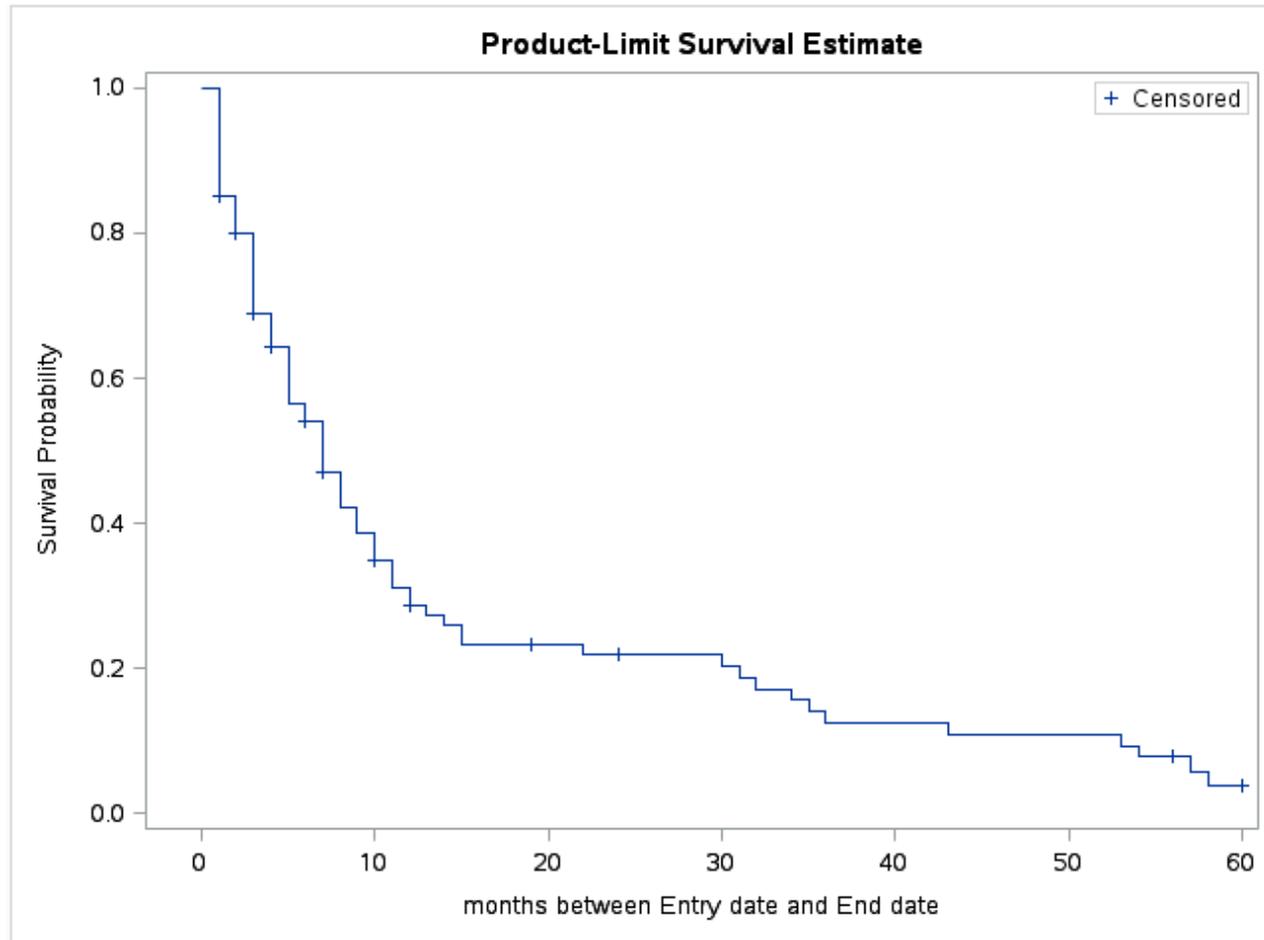
# Kaplan-Meier Estimate

## SAS code

```
libname present '/folders/myfolders/Presentation/';  
data hmohiv;  
    set present.hmohiv;  
run;  
ods listing close;  
ods output ProductLimitEstimates=est;  
proc lifetest data=hmohiv plots=(s);  
    time time*censor(0);  
run;  
ods listing;
```

# Kaplan-Meier Estimate

## SAS output



# Tests and Examples

TEST	EXAMPLE
One-sample t-test	Joint position sense
Two-sample t-test	A small clinical study: weight
Paired t-test	Study: bran in the treatment of diverticulosis
One-sample proportion test	Test the proportion of babies
Two-sample proportion test	Gender differences relative to smoking behavior
McNemar's test	Treatments for athlete's foot
Kaplan-Meier Estimate	HMO-HIV+ Study
<b>Log rank test</b>	<b>HMO-HIV+ Study</b>

# Log-rank test

- Test to compare the survival distributions of two or more samples
- Nonparametric test for right skewed and censored
- Widely used in clinical trials on the efficacy of a new treatment in comparison with a control treatment when the measurement is the time to event

# Log-rank test

Example: HMO-HIV+ study

- Tests of equality of the survivorship functions across the two drug strata
- The null hypothesis
  - No difference between survival curves
- The alternative hypothesis
  - The survival curves are different

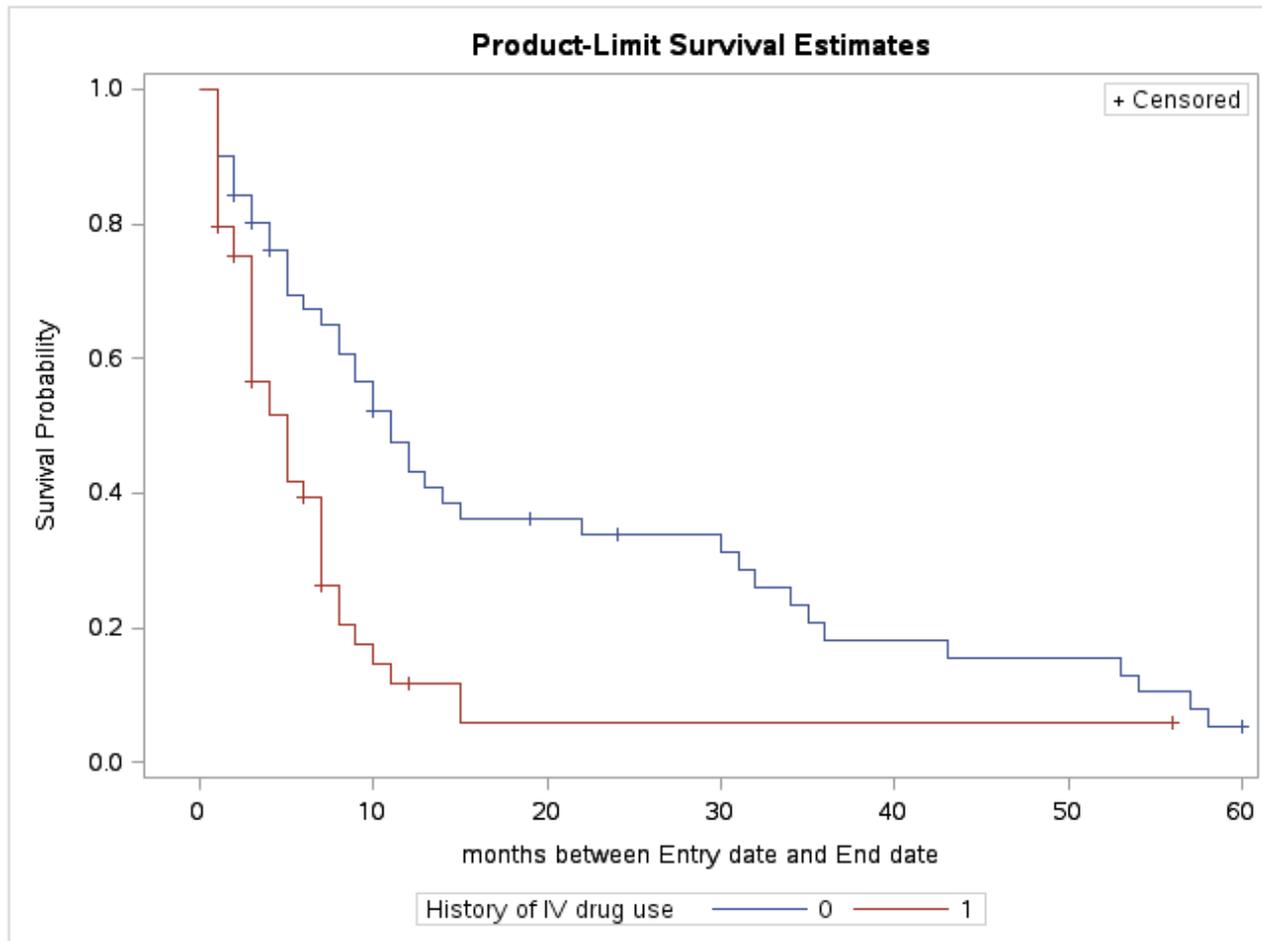
# Log-rank test

## SAS code

```
proc lifetest data=hmo hiv plots=(s);  
  time time*censor(0);  
  strata drug /test=(logrank wilcoxon tarone peto);  
run;
```

# Log rank test

## SAS output



# Log rank test

## SAS output

Test of Equality over Strata			
Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	11.8556	1	0.0006
Wilcoxon	10.9104	1	0.0010
Tarone	12.3359	1	0.0004
Peto	11.4974	1	0.0007

# Help is Available

- **CTSC Biostatistics Office Hours**
  - Every Tuesday from 12 – 1:30 in Sacramento
  - Sign-up through the CTSC Biostatistics Website
- **EHS Biostatistics Office Hours**
  - Every Monday from 2-4 in Davis
- **Request Biostatistics Consultations**
  - CTSC - [www.ucdmc.ucdavis.edu/ctsc/](http://www.ucdmc.ucdavis.edu/ctsc/)
  - MIND IDDRC - [www.ucdmc.ucdavis.edu/mindinstitute/centers/iddrc/cores/bbrd.html](http://www.ucdmc.ucdavis.edu/mindinstitute/centers/iddrc/cores/bbrd.html)
  - Cancer Center and EHS Center