

Selection of Appropriate Statistical Analysis



Presented by:

Mohammad Moqaddasi Amiri

Assistant Professor of Biostatistics

Sirjan Faculty of Medical Sciences

April 2021

Objectives

- Getting to know about the statistical tests
- Selection the appropriate test based on the goals of the study

Outline

- Introduction
 - Types of variables
 - Types of studies
- Statistical Tests
 - Continuous outcome
 - Parametric
 - Non-Parametric
 - Categorical outcome

Introduction

- Types of Variables
 - Quantitative (Continuous)
 - Qualitative (Categorical)
 - Binary
 - Nominal
 - Ordinal

Introduction

- Cut points
 - Convert the quantitative to the qualitative variables
 - Age groups
 - Questionnaire scores
 - Blood sugar
 - Hypertension

Introduction

- انواع مطالعه

- توصیفی

- آمار توصیفی (تعیین میانگین یا شیوع یا تعیین فاصله اطمینان)

- تحلیلی

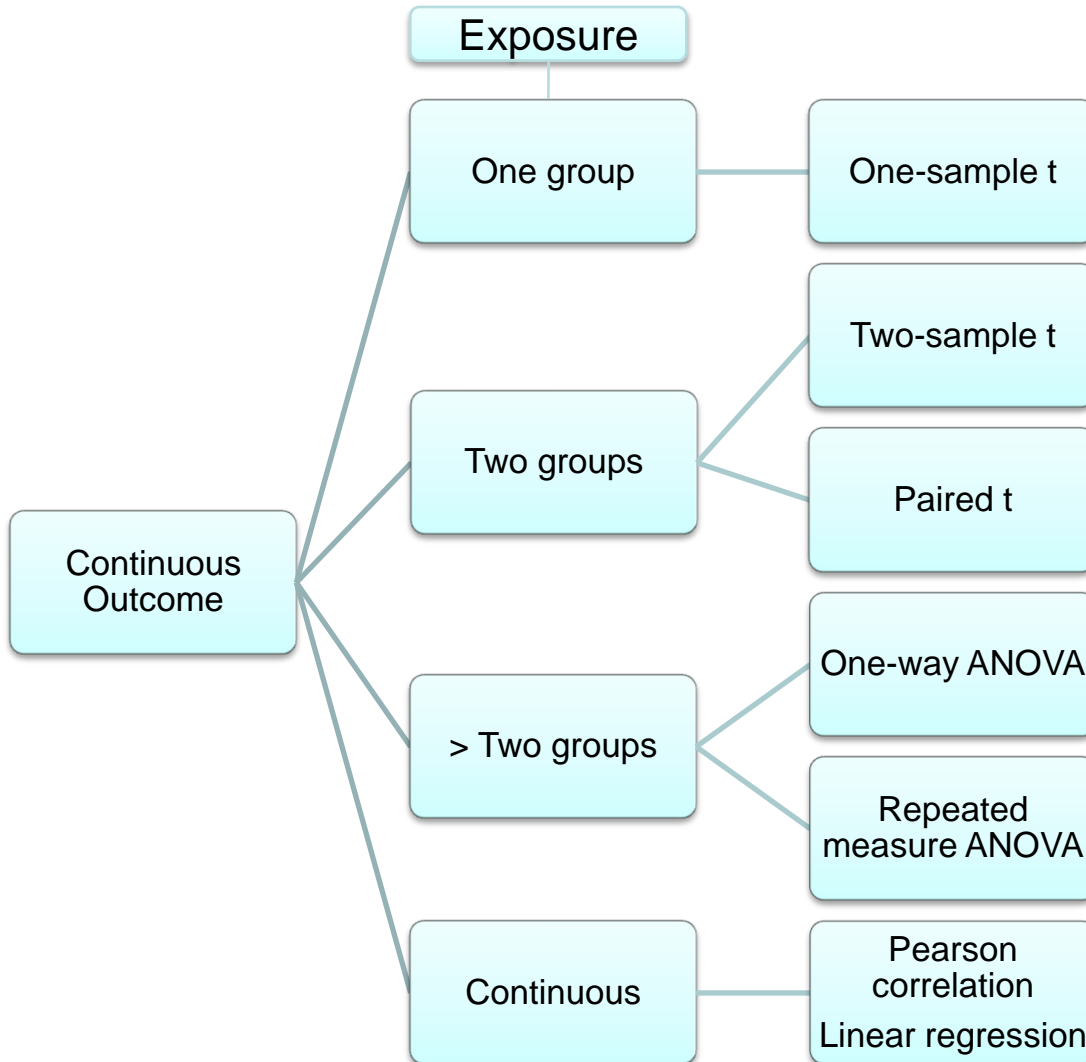
- مقایسه در گروه های مختلف

- تعیین ارتباط

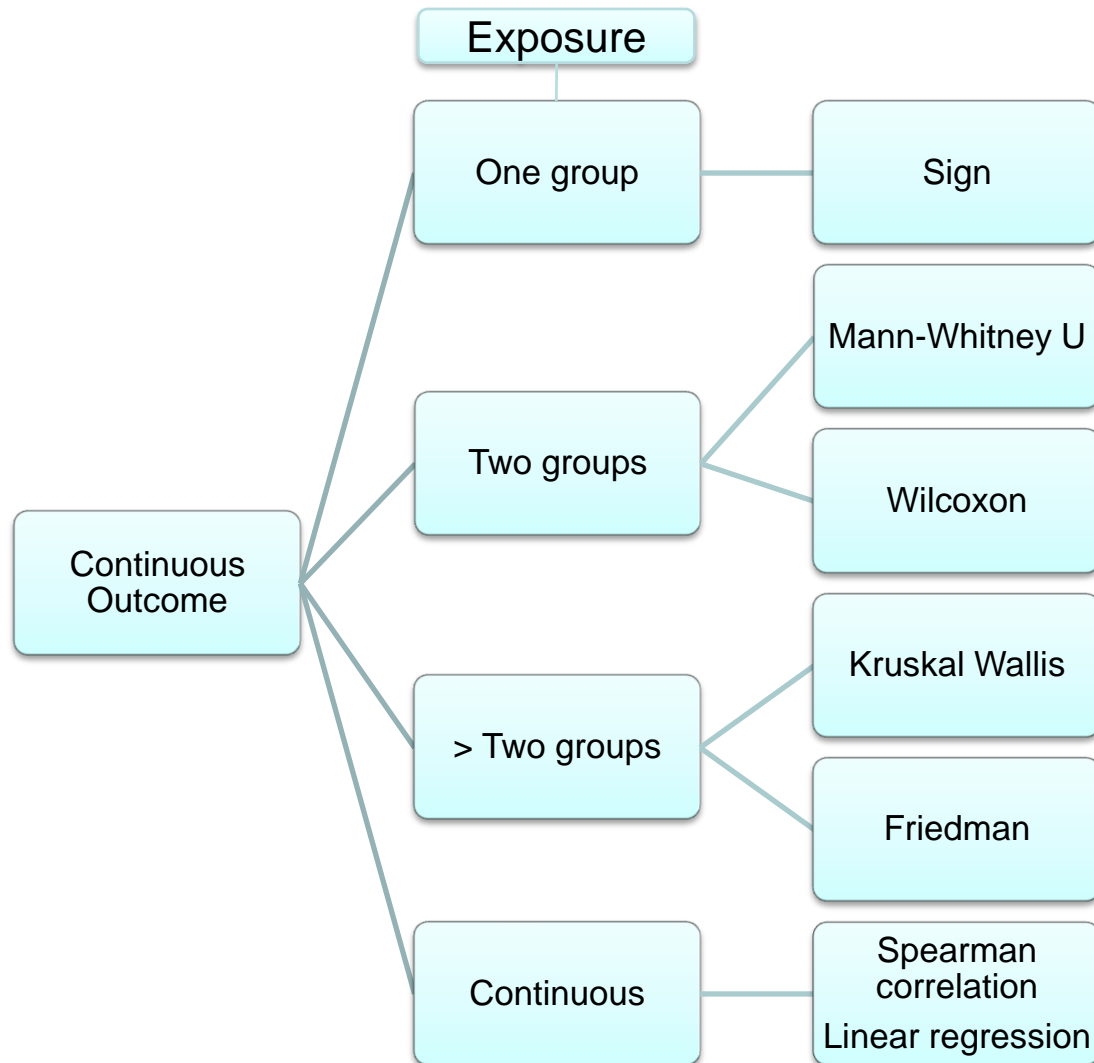
Statistical Analysis

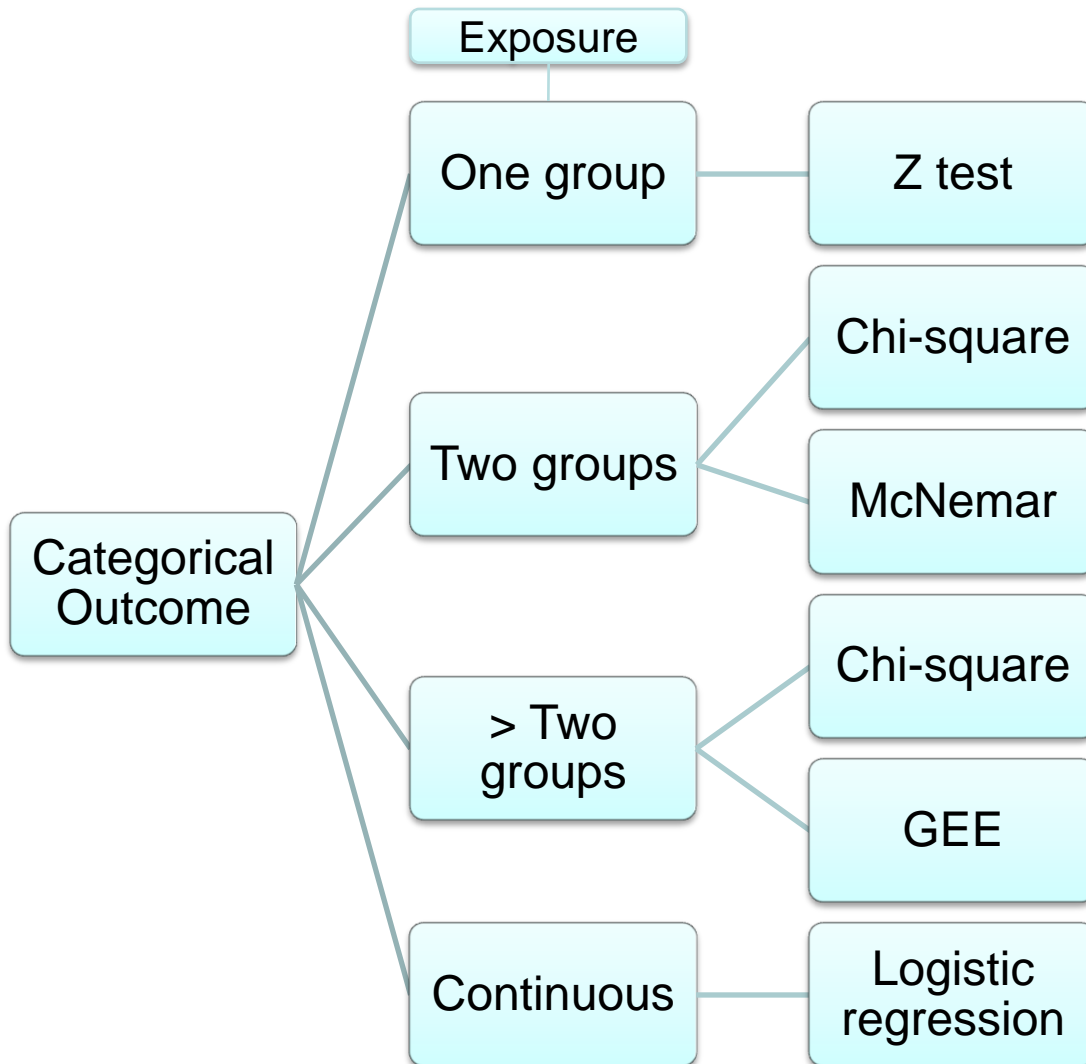
- Continuous outcome (response, dependent)
 - Parametric methods
 - Non-Parametric methods
- Categorical outcome

Parametric



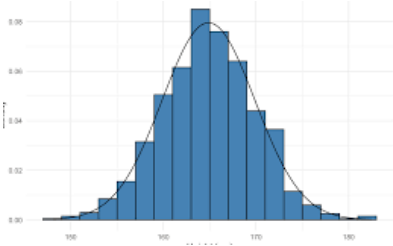
Non-Parametric



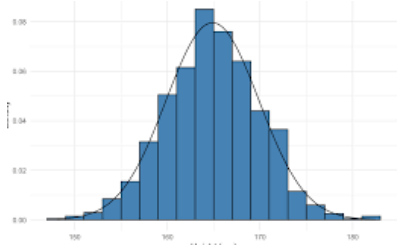


Continuous Outcome

- هدف: بررسی نرمال بودن متغیر کمی

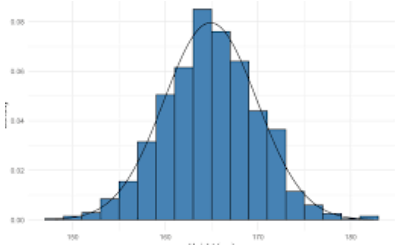


Continuous Outcome



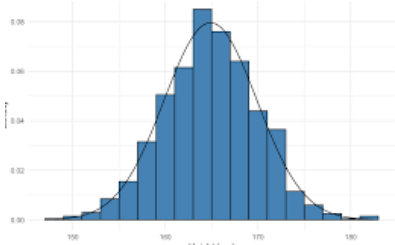
- هدف: بررسی نرمال بودن متغیر کمی
 - آزمون شاپیرو-ویلک یا کولموگروف-اسمیرنوف
 - نرمالیتی بایستی در تمامی زیرگروه ها بررسی شود

Continuous Outcome



- هدف: بررسی نرمال بودن متغیر کمی
 - آزمون شاپیرو-ویلک یا کولموگروف-اسمیرنوف
 - نرمالیتی بایستی در تمامی زیرگروه ها بررسی شود
 - روش های ترسیمی (هیستوگرام، شاخه و برگ، P-P plot و Q-Q plot)
 - برای حجم نمونه بالای ۲۰۰ معمولاً به بررسی نرمالیتی نیازی نیست.

Continuous Outcome



- **هدف: بررسی نرمال بودن متغیر کمی**

- آزمون شاپیرو-ویلک یا کولموگروف-اسمیرنوف
- نرمالیتی بایستی در **تمامی** زیرگروه ها بررسی شود
- روش های ترسیمی (هیستوگرام، شاخه و برگ، P-P plot و Q-Q plot)
- برای حجم نمونه بالای **۲۰۰** معمولاً به بررسی نرمالیتی نیازی نیست.

- **مثال:**

- بررسی نرمال بودن قد دانشجویان
- بررسی نرمال بودن منیزیم سرم در مردان و زنان
- بررسی نرمال بودن میزان کلسیم در گروه های سنی مختلف

Continuous Outcome

- هدف: مقایسه میانگین یک متغیر کمی با یک مقدار ثابت
 - آزمون t یک نمونه ای (**One-Sample t test**)
 - یک متغیر کمی داریم و یک جامعه (گروه)



Continuous Outcome

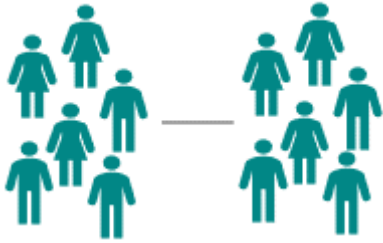
- هدف: مقایسه میانگین یک متغیر کمی با یک مقدار ثابت
 - آزمون t یک نمونه ای (One-Sample t test)
 - یک متغیر کمی داریم و یک جامعه (گروه)



- مثال: در مطالعات گذشته میانگین وزن نوزادان متولد شده از مادران فقیر ۲۸۰۰ گرم بوده است. پژوهشگری با یک برنامه غذایی قصد دارد تا وزن نوزادان متولد شده در خانواده های فقیر را افزایش دهد.

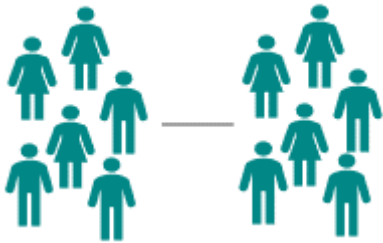
Continuous Outcome

- هدف: مقایسه میانگین یک متغیر کمی در دو جامعه
 - آزمون t دو نمونه ای مستقل (**Two-Sample t test**)
 - **یک** متغیر کمی داریم و **دو** جامعه (یک متغیر کیفی دو حالت)



Continuous Outcome

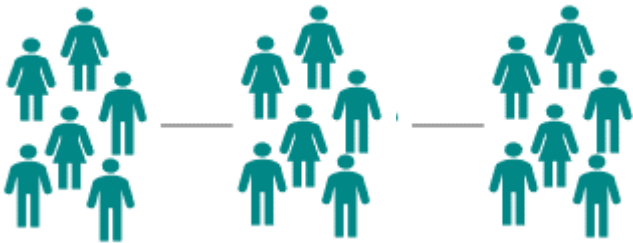
- هدف: مقایسه میانگین یک متغیر کمی در دو جامعه
 - آزمون t دو نمونه ای مستقل (**Two-Sample t test**)
 - **یک** متغیر کمی داریم و **دو** جامعه (یک متغیر کیفی دو حالته)



- مثال:
- مقایسه میزان قند خون در مردان و زنان
- مقایسه نمره سلامت اجتماعی افراد مجرد و متأهل
- بررسی اثر داروی فشار خون در گروه های مورد و شاهد

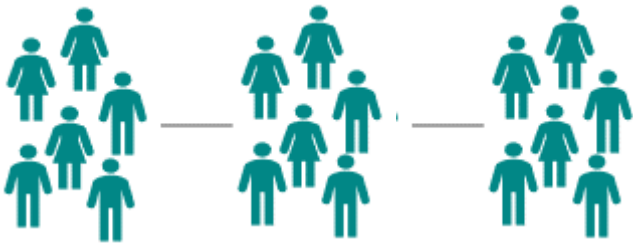
Continuous Outcome

- هدف: مقایسه میانگین یک متغیر کمی در بیش از دو جامعه
 - روش آنالیز واریانس یک طرفه (One-Way ANOVA)
 - یک متغیر کمی داریم و بیش از دو جامعه (یک متغیر کیفی چندحالتی)



Continuous Outcome

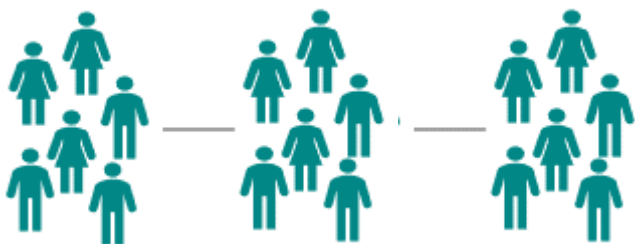
- هدف: مقایسه میانگین یک متغیر کمی در بیش از دو جامعه
 - روش آنالیز واریانس یک طرفه (One-Way ANOVA)
 - یک متغیر کمی داریم و بیش از دو جامعه (یک متغیر کیفی چندحالتی)



- به غیر از نرمالیتی نیاز به بررسی همگنی واریانس ها می باشد.
- آزمون Levene
- اگر تعداد نمونه در زیرگروه ها تقریباً برابر باشد نیاز به بررسی این فرضیه نیست

Continuous Outcome

- هدف: مقایسه میانگین یک متغیر کمی در بیش از دو جامعه
 - روش آنالیز واریانس یک طرفه (One-Way ANOVA)
 - یک متغیر کمی داریم و بیش از دو جامعه (یک متغیر کیفی چندحالتی)



- مثال:
- بررسی میزان قند خون در گروه های سنی مختلف
- مقایسه میزان کلسترول در سه گروه درمانی
- آیا میزان سلامت روان افراد در گروه های درآمدی مختلف با هم متفاوت است؟

Continuous Outcome

- هدف: مقایسه میانگین یک متغیر کمی در جفت های همسان (Matched pairs)

– آزمون t زوجی (Paired t test)

– یک متغیر کمی به صورت قبل و بعد برای هر فرد (before-after)



Continuous Outcome

- هدف: مقایسه میانگین یک متغیر کمی در جفت های همسان (Matched pairs)

– آزمون t زوجی (Paired t test)

– یک متغیر کمی به صورت قبل و بعد برای هر فرد (before-after)



– مثال:

- مقایسه میزان قند خون قبل و بعد از مصرف دارو
- مقایسه فشار خون افراد در دو حالت نشسته و خوابیده
- مقایسه نمره سلامت اجتماعی افراد قبل و بعد از یک مداخله آموزشی

Continuous Outcome

- هدف: مقایسه میانگین یک متغیر کمی در جفت های همسان
(Matched pairs)

– آزمون t زوجی (Paired t test)

– یک متغیر کمی به صورت قبل و بعد برای هر فرد (before-after)

– بررسی یک متغیر کمی در دو گروه با همسان سازی فردی
(individual matching)













Matching

- Individual matching
- Group (Frequency) matching

Matching

- Individual matching













by: age, sex and race

Cases		Controls	
 40	 45	 40	 45
 50	 55	 50	 55
 60	 65	 60	 65

Matching

- Group matching

by: age, sex and race

Cases		Controls	
 45	 40	 40	 45
 55	 50	 50	 55
 65	 60	 60	 65

Continuous Outcome

- هدف: مقایسه میانگین یک متغیر کمی در جفت های همسان (Matched pairs)

– آزمون t زوجی (Paired t test)

– یک متغیر کمی به صورت قبل و بعد برای هر فرد (before-after)

– بررسی یک متغیر کمی در دو گروه با همسان سازی فردی (individual matching)

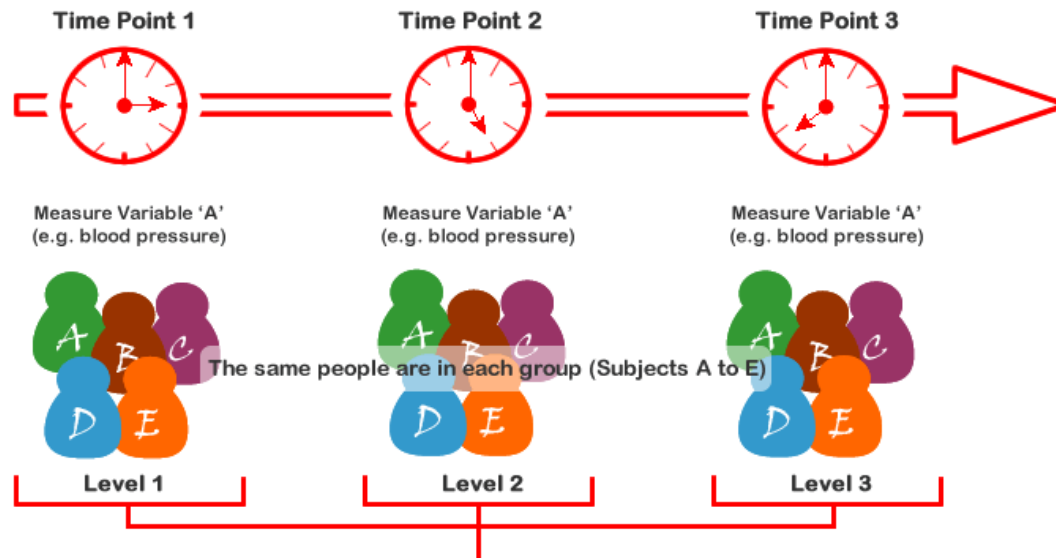
– مثال:

– مقایسه میزان قند خون در مردان و زنان که بر اساس سن همسان شده اند

– مقایسه نمره سلامت اجتماعی افراد مجرد و متأهل همسان شده بر اساس سن و جنسیت

Continuous Outcome

- هدف: مقایسه میانگین یک متغیر کمی در زمان های مختلف یا در گروه های همسان مختلف
- آزمون آنالیز واریانس اندازه های تکراری (Repeated measure ANOVA)
- یک متغیر کمی که در زمان های مختلف اندازه گیری می شود



Continuous Outcome

- هدف: مقایسه میانگین یک متغیر کمی در زمان های مختلف یا در گروه های همسان مختلف
- آزمون آنالیز واریانس اندازه های تکراری (Repeated measure ANOVA)
- **یک** متغیر کمی که در زمان های مختلف اندازه گیری می شود

مثال:

بررسی سطح اضطراب در دوره های ۳ ماهه
بررسی فشار خون افراد در حالت های ایستاده، نشسته و خوابیده
بررسی تاثیر یک دارو بر روی مقدار قند خون هر دو ماه یکبار

Continuous Outcome

- هدف: مقایسه میانگین یک متغیر کمی در زمان های مختلف یا در گروه های همسان مختلف
 - آزمون آنالیز واریانس اندازه های تکراری (Repeated measure ANOVA)
 - **یک** متغیر کمی که در زمان های مختلف اندازه گیری می شود
 - مقایسه یک متغیر کمی در بیش از دو گروه با **individual matching**

Continuous Outcome

- هدف: مقایسه میانگین یک متغیر کمی در زمان های مختلف یا در گروه های همسان مختلف
 - آزمون آنالیز واریانس اندازه های تکراری (Repeated measure ANOVA)
 - **یک** متغیر کمی که در زمان های مختلف اندازه گیری می شود
 - مقایسه یک متغیر کمی در بیش از دو گروه با **individual matching**

مثال:

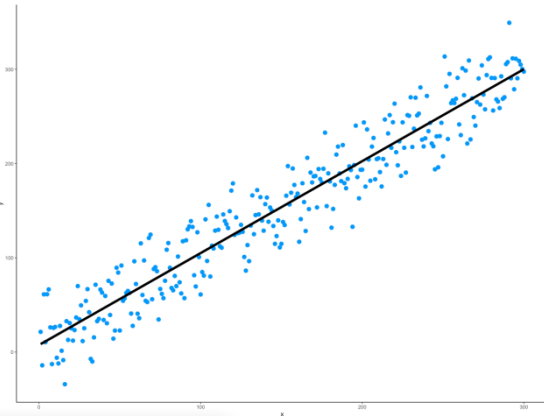
مقایسه شاخص توده بدنی در گروه های درآمدی مختلف که بر اساس سن و جنس همسان سازی شده اند.

بررسی میزان قند خون در شغل های مختلف که بر اساس BMI و جنسیت همسان شده اند.

Continuous Outcome

- هدف: ارتباط یک متغیر کمی با یک متغیر کمی دیگر

– ضریب همبستگی پیرسن (Pearson correlation coefficient) یا رگرسیون خطی (linear regression)

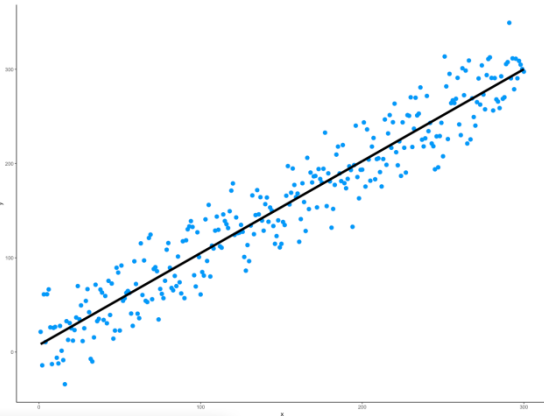


– دو متغیر کمی

Continuous Outcome

- هدف: ارتباط یک متغیر کمی با یک متغیر کمی دیگر

– ضریب همبستگی پیرسن (Pearson correlation coefficient) یا رگرسیون خطی (linear regression)





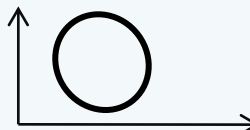
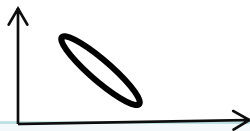

– دو متغیر کمی

- مثال:
- بررسی ارتباط بین سن و میزان آگاهی اجتماعی
- بررسی ارتباط بین کلسترول خون و تری گلیسرید

Correlation

- Value between -1 and 1
 - Direct relationship ($r > 0$)
 - Inverse relationship ($r < 0$)
 - No relationship ($r = 0$)

Correlation

ضریب همبستگی	نمودار پراکنش	تفسیر
$\rho_{xy} = 1$		همبستگی دو متغیر کامل و مستقیم است
$0 < \rho_{xy} < 1$		همبستگی دو متغیر ناقص و مستقیم است
$\rho_{xy} = 0$		دو متغیر همبستگی ندارند. (ناهمبسته)
$-1 < \rho_{xy} < 0$		همبستگی دو متغیر ناقص و معکوس است
$\rho_{xy} = -1$		همبستگی دو متغیر کامل و معکوس است

Correlation

Size of Correlation	Interpretation
.90 to 1.00 (-.90 to -1.00)	Very high positive (negative) correlation
.70 to .90 (-.70 to -.90)	High positive (negative) correlation
.50 to .70 (-.50 to -.70)	Moderate positive (negative) correlation
.30 to .50 (-.30 to -.50)	Low positive (negative) correlation
.00 to .30 (.00 to -.30)	negligible correlation

Continuous Outcome

- **One-sample t-test**
- **Two-independent sample t test**
 - normality
- **One-way ANOVA**
 - normality
 - Homogeneity of variances
- **Matched paired t-test**
 - before-after
 - Individual matching of two groups
- **Repeated measures ANOVA**
 - measurement in several time points or situation
 - Individual matching of more than two groups
- **Pearson correlation / Linear regression**

Continuous Outcome

- Non-Parametric tests
 - abnormal variable
 - Absence of homogeneity of variances

Continuous Outcome

Situation of exposure	Parametric tests	Non-Parametric tests
One group	One-sample t test	Sign test
Two groups	Two-independent sample t-test	Mann-Whitney U
Two groups with individual matching / before-after	Matched paired t-test	Wilcoxon
More than two groups	One-way ANOVA	Kruskal-Wallis
Measurement in several time points / Individual matching of more than two groups	Repeated measures ANOVA	Friedman
Continuous	Pearson correlation / Linear regression	Spearman correlation / Linear regression

Categorical Outcome

- هدف: مقایسه شیوع (نسبت) یک متغیر کیفی با یک مقدار ثابت
– آزمون Z یک نمونه ای (One-Sample Z test)
– یک متغیر کیفی داریم و یک جامعه (گروه)



Categorical Outcome

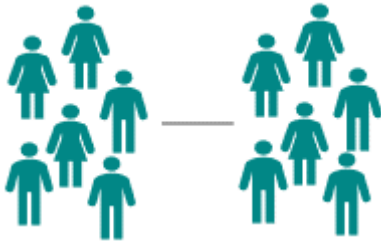
- هدف: مقایسه شیوع (نسبت) یک متغیر کیفی با یک مقدار ثابت
 - آزمون Z یک نمونه ای (One-Sample Z test)
 - یک متغیر کیفی داریم و یک جامعه (گروه)



- مثال: در مطالعات گذشته شیوع یک بیماری $0/3$ گزارش شده است. پروهشگری مدعی است که شیوع این بیماری بیش از $0/3$ است.

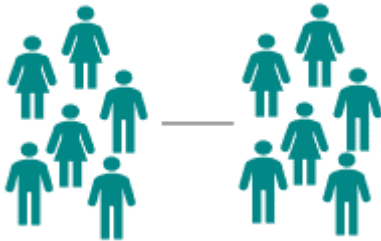
Categorical Outcome

- هدف: مقایسه شیوع یک متغیر کیفی در دو جامعه
– آزمون مجذور کای (**Chi-square test**)
– **یک** متغیر کیفی داریم و یک متغیر کیفی دو حالتی



Categorical Outcome

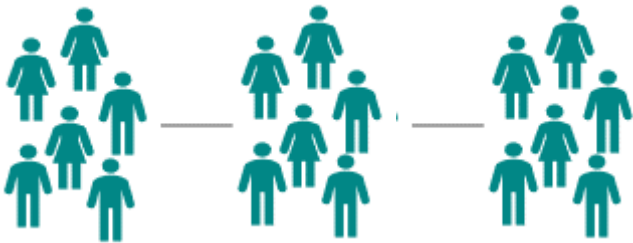
- هدف: مقایسه شیوع یک متغیر کیفی در دو جامعه
 - آزمون مجذور کای (**Chi-square test**)
 - **یک** متغیر کیفی داریم و یک متغیر کیفی دو حالتی



- مثال:
- مقایسه شیوع دیابت در مردان و زنان
- مقایسه نسبت افراد با سلامت معنوی بالا و پایین در افراد مجرد و متأهل
- بررسی شیوع فشار خون در دو گروه سنی

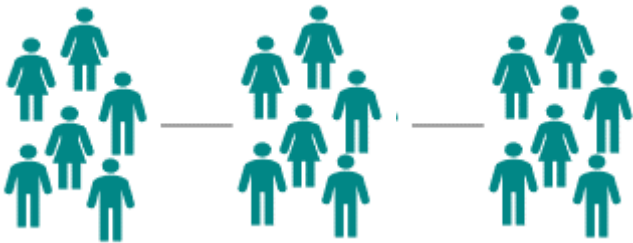
Categorical Outcome

- هدف: مقایسه شیوع یک متغیر کیفی در بیش از دو جامعه
– آزمون مجذور کای (**Chi-square test**)
– **یک** متغیر کیفی داریم و یک متغیر کیفی چندحالتی



Categorical Outcome

- هدف: مقایسه شیوع یک متغیر کیفی در بیش از دو جامعه
 - آزمون مجذور کای (**Chi-square test**)
 - **یک** متغیر کیفی داریم و یک متغیر کیفی چندحالتی



- مثال:
- بررسی شیوع دیابت در گروه های سنی مختلف
- بررسی شیوع کلسترول بالا در سه گروه درمانی
- آیا میزان سلامت معنوی افراد (کم، زیاد) در گروه های درآمدی مختلف با هم متفاوت است؟

Categorical Outcome

- هدف: مقایسه شیوع یک متغیر کیفی در جفت های همسان (Matched pairs)

– آزمون **McNemar**

– یک متغیر کیفی به صورت قبل و بعد برای هر فرد (before-after)

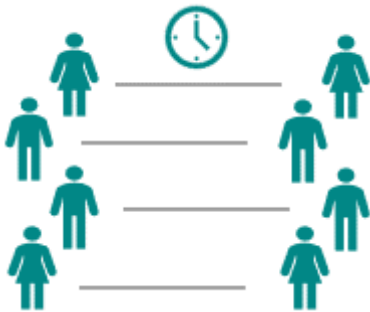


Categorical Outcome

- هدف: مقایسه شیوع یک متغیر کیفی در جفت های همسان (Matched pairs)

– آزمون McNemar

– یک متغیر کیفی به صورت قبل و بعد برای هر فرد (before-after)



– مثال:

- مقایسه سلامت اجتماعی افراد (کم و زیاد) قبل و بعد از یک مداخله آموزشی
- بررسی نسبت افراد با فشار خون بالا قبل و بعد از یک دوره درمانی

Categorical Outcome

- هدف: مقایسه شیوع یک متغیر کیفی در جفت های همسان (Matched pairs)

– آزمون McNemar

– یک متغیر کیفی به صورت قبل و بعد برای هر فرد (before-after)

– بررسی یک متغیر کیفی در دو گروه با همسان سازی فردی (individual matching)

Categorical Outcome

- هدف: مقایسه شیوع یک متغیر کیفی در جفت های همسان (Matched pairs)

– آزمون McNemar

– یک متغیر کیفی به صورت قبل و بعد برای هر فرد (before-after)

– بررسی یک متغیر کیفی در دو گروه با همسان سازی فردی (individual matching)

– مثال:

- بررسی شیوع دیابت در دو گروه مردان و زنان که بر اساس سن و BMI همسان شده اند.
- بررسی سطح سلامت روان در افراد متأهل و مجرد که بر اساس جنسیت و سن همسان شده اند.

Categorical Outcome

- هدف: مقایسه شیوع یک متغیر کیفی در زمان های مختلف و یا در گروه های مختلف با **individual matching**
- آزمون **Generalized Estimation Equation (GEE)**
- **یک** متغیر کیفی که در زمان های مختلف اندازه گیری می شود
- مقایسه یک متغیر کیفی در گروه های با **individual matching**

Categorical Outcome

- هدف: مقایسه شیوع یک متغیر کیفی در زمان های مختلف و یا در گروه های مختلف با **individual matching**
- آزمون **Generalized Estimation Equation (GEE)**
 - **یک** متغیر کیفی که در زمان های مختلف اندازه گیری می شود
 - مقایسه یک متغیر کیفی در گروه های با **individual matching**
- مثال:
- بررسی سطح اضطراب (کم، زیاد) در دوره های ۳ ماهه
- مقایسه چاقی در گروه های درآمدی مختلف که بر اساس سن و جنس همسان شده اند.

Categorical Outcome

- هدف: ارتباط یک متغیر کیفی با یک متغیر کمی

– رگرسیون لجستیک (Logistic regression)

Categorical Outcome

- هدف: ارتباط یک متغیر کیفی با یک متغیر کمی

– رگرسیون لجستیک (Logistic regression)

– مثال:

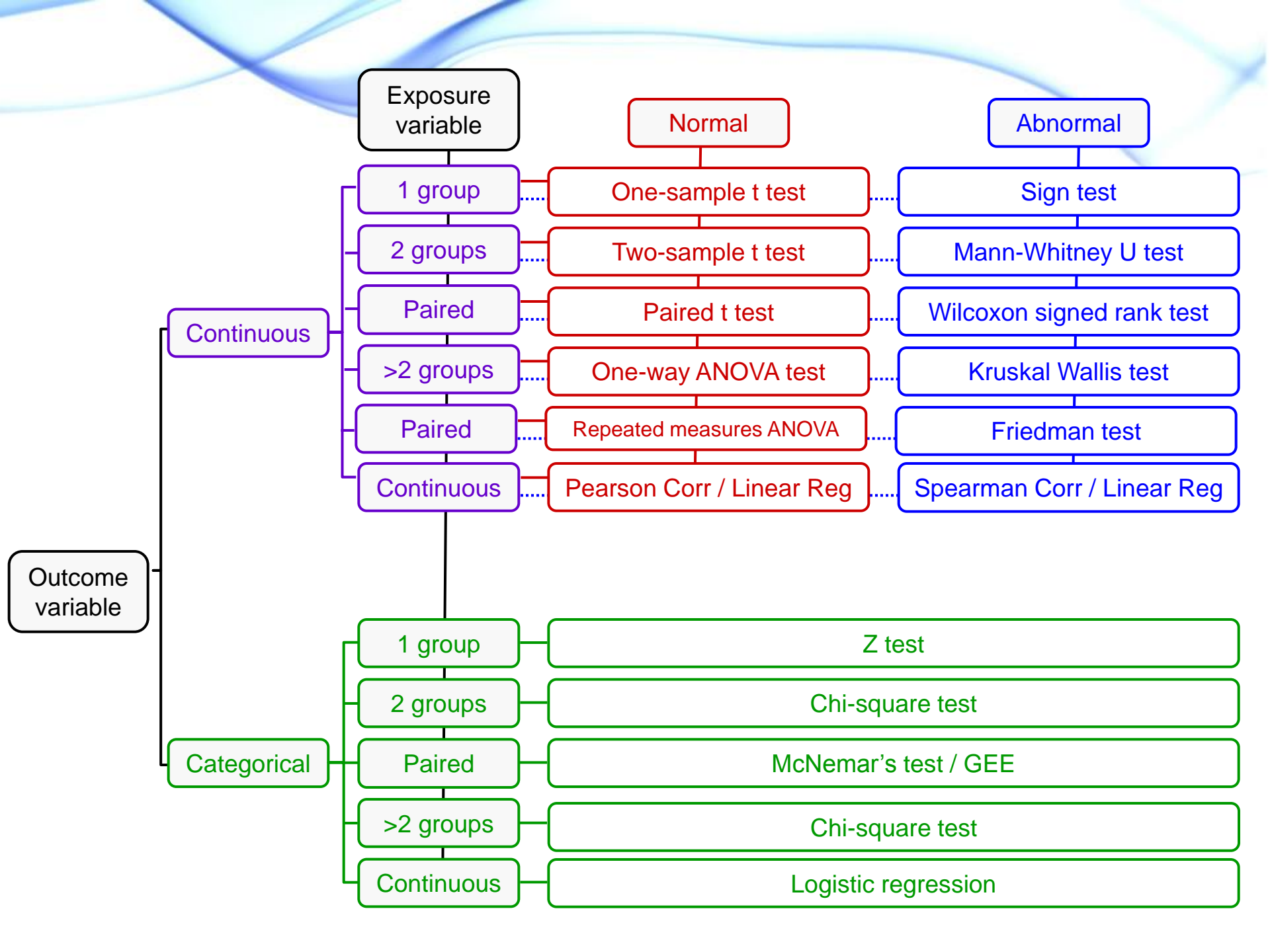
- بررسی تأثیر سن بر آگاهی اجتماعی (کم و زیاد)
- بررسی ارتباط بین شاخص توده بدنی و شیوع فشار خون بالا

Categorical Outcome

- **One sample Z test**
- **Chi-square test**
- **McNemar's test**
- **GEE**
- **Logistic regression**

Advanced Methods

- Regression Analysis
 - Multiple linear / logistic regression / ANCOVA
 - ordinal / nominal / poisson / negative-binomial regression
- Survival Analysis
 - Log-rank test / Cox regression
- Longitudinal Data Analysis
 - (Generalized)Linear Mixed-effects Models (LMM & GLMM)
- Multivariate Data Analysis
 - MANOVA / multivariate regression



Thank You