BIOSTATISTICS

Definitions

Parametric Tests	Non-Parametric Tests (Distribution-free)
Assume population is normally, homogenously, & independently distributed.	Don’t require this assumption. Observation is independent.
- x’ - SD - *t*-test - ANOVA - Pearson Coefficient -Kolmogorov Smirnov test (Fischer’s exact test & Mantel-Haenszel test are extensions of χ^2)	- Median - Mode - χ^2 - Spearman Coefficient - Mann-Whitney U test - Wilcoxon’s signed rank test - Kruskal Wallis test - Friedman test.

Accuracy	Closure of value to true value.
Precision (Kappa; κ = 0-1) κ = 0 → no agreement, >0 & <1 → good agreement, 1 → perfect agreement	Reproducibility of result (i.e, every time you repeat the test, it gives the same result)
If true value is unknown → precision more imp than accuracy	
Bias	Systematic difference from true value.
Statistical Errors = failure of statistical test	
Type I error (α) **= Probability (*P* value)**	Rejection of correct hypothesis (false +ve).
Chance of type I error = *P* value	
Type II error (β)	Acceptance of incorrect hypothesis (false -ve).
Depends on statistical power of study. ↑ statistical power → ↓ Type II error Power of study = 1 – β <u>Power depends on:</u> - Significance level. - Sample size. - Accuracy of measurements.	

Glossary p. 98 (Cont. variables) FPGEE Secrets p. 73 (independent vs. depend…)
Glossary p. 99 (Discrete “) FPGEE Secrets p. 72, 74

Important Statistical Values & Tests

n: no. *P*: probability

Name	Symbol	Equation/Definition
Mean	**x’**	$\frac{\Sigma x}{n}$ **also,** $n \times P$
Median	-	Mean of the two middle values
Mode	-	Most common occurring value
Frequency of Distribution	-	Mean – Median (Mode) In graph: If result: +ve → +ve skewed -ve → -ve skewed 0 → normal distribution
Range of Set	-	Highest value – Lowest value
Binomial Distribution	-	$n \times P$
Standard Deviation	**SD**	$\sqrt{\frac{\Sigma(x - x')^2}{n - 1}}$
x' ± SD		
Degree of Freedom	***df***	$n - 1$
Degree of Freedom in Chi-square table	***df***	$(R - 1) \times (C - 1)$ R: no. of rows C: no. of columns
Variance	-	$(SD)^2$
Relative Standard Deviation	**%RSD**	$\frac{SD}{x'} \times 100$
Standard Deviation of Mean	**S_m**	$\frac{SD}{\sqrt{n}}$
Pearson Coefficient	**r**	$\frac{3\,(\text{Mean} - \text{Median})}{SD}$

With SD
↓
Precision

Chi-square **(Cross breaks)** To improve accuracy of *P* value, apply "Yates' continuity correction"	χ^2	$\Sigma \frac{([E-O]-0.5)^2}{E}$ E: Expected O: Observed If χ^2 = zero → null hypothesis ↑ χ^2 → more significant
Risk Ratio (Relative Risk) In Cohort studies	-	**Risk =** $\frac{\text{no.of events}}{\text{no.of ppl exposed to that event}}$ **Risk ratio =** $\frac{\text{risk in tratment (exposed)gp}}{\text{risk in control (unexposed)gp}}$ Results: 1 = no risk > 1 = exposure ↑ risk < 1 = exposure ↓ risk *If CI (confidence interval) of risk ratio includes 1 → statistically insignificant (& vice versa)
Odds Ratio In case-control studies	-	**Odd =** $\frac{\text{no.of times the event happen}}{\text{no.of times the event } \textbf{not} \text{ happen}}$ **Odd ratio =** $\frac{\text{odds of being exposed to risk factor}}{\text{odds in control gp}}$ Results: 1 = no diff in risk b/w gps > 1 = risk of event ↑ in exposure < 1 = risk of event ↓ in exposure *If CI (confidence interval) of odd ratio includes 1 → statistically insignificant (& vice versa)
Absolute Risk Reduction	**ARR**	**ARR=** improvement (event) rate in ttt gp (%) – improvement (event) rate in control gp (%)
Number Needed to Treat	**NNT**	$\mathbf{NNT} = \frac{100}{\text{ARR}}$ i.e, NNT of pts should be tted for **<u>1</u>** to get benefit. (↓ NNT → the better)
Relative Risk Reduction	**RRR**	% of ↓ of risk (disease) from control gp to ttt gp **RRR=** $\frac{\text{\% of unimproved (diseased) pts in ctrl gp} - \text{\% of unimproved pts in ttt gp}}{\text{\% of unimproved pts in ttt gp}} \times 100$
Number Needed to Harm	**NNH**	**NNH=** $\frac{100}{\text{\% of pts had SE in ttt gp} - \text{\% of pts had SE in ctrl gp}}$

Hazard Ratio (Cox Regression Model; Proportional Hazards Survival Model) -Estimate of life expectancy -Describes relationship b/w event (usually death) & variables (e.g. smoking).	**HR**	$\textbf{HR}= \frac{\text{hazard of event in gp 1}}{\text{hazard of event in gp 2}}$ Hazard: chance of something harmful happening If HR= 1 → risk is same b/w 2 gps If HR= 2 → risk is double in gp 1 than gp 2
Correlation Coefficient +ve R^2 = as one variable ↑ the other variable is also ↑ -ve R^2 = as one variable ↓ the other variable is also ↓	R^2	$\mathbf{R^2} = (r)^2$ → × 100 = % i.e, % of variation in (y) axis is related to variation in (x) axis * The closest to **<u>1</u>** → the strongest the correlation, whether +ve or –ve (r)
*Pearson Correlation Coefficient is used if **normal** distribution, otherwise Spearman Correlation Coefficient is used.		
Regression	-	$y = a + bx$ b: regression coefficient

Types of regression:

Logistic regression:

Used where each case in the sample can only belong to one of two groups (e.g. having disease or not) with the outcome as the probability that a case belongs to one group only .

Poisson regression:

Used to study waiting times or time between rare events.

Difference between Correlation and Regression:

- Correlation measures the **strength** of the association b/w variables.
- Regression **quantifies** the association.

Kaplan-Meier:

Survival analysis (life tables).

Cox proportional hazards regression model:

Used in survival analysis where the outcome is time until a certain event.

Sensitivity, Specificity, & Predictive value:

		Disease	
		Present	**Absent**
Test result	**Positive**	A	B (false +ve)
	Negative	C (false -ve)	D

	Equation	Definition
All diseased	**Sensitivity =** $\frac{A}{A+C}$	How often the test will be +ve if the pt really have the disease
All healthy	**Specificity =** $\frac{D}{D+B}$	How often the test will be -ve if the pt is really healthy
All +ve	**Positive Predictive Value (PPV) =** $\frac{A}{A+B}$	If the result is +ve, what is the likelihood that the pt really have the disease
All -ve	**Negative Predictive Value (NPV) =** $\frac{D}{D+C}$	If the result is -ve, what is the likelihood that the pt is really healthy

- **Perfect test if all = 1**
- ↓ the value → the test less useful
- All × 100 = %

Likelihood ratio (LR): **LR =** $\frac{\text{Sensitivity}}{1-\text{Specificiy}}$

if the test is +ve, how much more likely the pt is to have the disease than not having it.

Pharmacy Management MCQs Book:
Questions: 250, 264, 307, 324