

Tutorial 11 for STAT 3004

Final Review

WANG, Lijun

November 28, 2018

Department of Statistics, CUHK

Quick Review

Standard tests

Nature	Data	Conditions for validity	R function
Parametric tests			
Mean	1 sample 2 samples 2 samples 2 paired samples	$n > 30$ or normality Normality and equal variances Normality $n > 30$ or normality	<code>t.test(x,...)</code> <code>t.test(x,y,...)</code> <code>t.test(x,y,var.equal=F)</code> <code>t.test(x,y,paired=T)</code>
Variance	1 sample 2 samples 2 samples	Normality Normality Large sample size	<code>sigma2.test(x,...)</code> <code>var.test(x,y,...)</code> <code>asypm.test(x,y,...)</code>
Correlation	1 sample 2 samples	Normality, $H_0 : \rho = \rho_0$ Normality	<code>cor.test(x,y,..)</code> <code>cor.test.2.sample(x,y,...)</code>
Proportion	1 sample 1 sample 2 samples	$np \geq 5$ and $n(1 - p) \geq 5$ Large sample size	<code>prop.test(x,...)</code> <code>binom.test(x,...)</code> <code>prop.test(x,y,...)</code>
Independence tests			
χ^2 for independence Yates' χ^2 Fisher's exact test	Contingency table 2×2 table Contingency table	Theoretical counts ≥ 5 Theoretical counts ≥ 2.5	<code>chisq.test(.,correct=F)</code> <code>chisq.test()</code> <code>fisher.test()</code>
Tests of fit to a distribution			
Shapiro–Wilk χ^2 of fit to a distribution Kolmogorov–Smirnov	1 sample 1 sample 1 sample 2 samples	Theoretical counts ≥ 5	<code>shapiro.test(x,...)</code> <code>chisq.test()</code> <code>ks.test(x,..)</code> <code>ks.test(x,y)</code>
Tests of position			
Median Sign test Mann–Whitney Mann–Whitney Wilcoxon	1 sample 2 samples 2 paired samples 2 samples 2 samples 2 paired samples	$\min(n_1, n_2) \geq 10$ $\min(n_1, n_2) \leq 10$	<code>binom.test(x,)</code> <code>fisher.test(x,y,)</code> <code>binom.test(x,y,paired=T)</code> <code>wilcox.test(x,y,exact=F)</code> <code>wilcox.test(x,y)</code> <code>wilcox.test(x,y,paired=T)</code>

Examples

Cow study

The quantity of bacteria per cm^3 of milk from eight different cows is estimated after milking and 24h later. We wish to test whether the quantity of bacteria significantly increases with time.

Cow	Just after milking	24 h after milking
1	12,000	11,000
2	13,000	20,000
3	21,500	31,000
4	17,000	28,000
5	15,000	26,000
6	22,000	30,000
7	11,000	16,000
8	21,000	29,000

- Answer the question under the assumption of normality of the data.
- Answer the question using a sign test.
- Answer the question using a Wilcoxon's signed rank test.

Number of patients in the emergency ward

To study the variation of the number of emergency cases in a hospital, the number of patients was counted for the months of June, July and August. The results are

	June	July	August
Number of patients	1,500	1,600	1,450
Number of emergency patients	675	720	610

Can we conclude that the proportion of emergency cases is the same every month?

Cell phone and driving reaction time

Let's assume that you know from past experience that reaction time has a standard deviation of 1.25 seconds. Also suppose that a 1-second difference in reaction time is considered an important difference. You want to conduct a two-sided significance test at 5% level with 90% power to detect such a difference if it exists. How many participants will you need in your study?

To design such experiment, what else do you need?

Gook luck!