Non-parametric Tests for 2 Independent Samples

- Test $H_o: \tilde{\mu}_1 = \tilde{\mu}_2$ as opposed to $H_o: \mu_1 = \mu_2$
- Use recoded data: the ranks of the observations as opposed to the original values.
- Do not rely on the assumption of normally distributed populations → Nonparametric

(version 1) A *T*-test on the Ranks

- Rank the combined data from both samples (ties get assigned the average rank).
- Do a 2-sample *t*-test on the ranks, instead of the observed values.
  - This test is roughly equivalent to the Wilcoxon Rank Sum Test.

(version 2) Wilcoxon Rank Sum Test (a.k.a. Mann-Whitney Test)

- Rank the combined data from both samples (ties get assigned the average rank).
- S = Sum of the ranks in sample 1 (called *W* in some software).

When H_o is true, the sampling distribution for S is approx. normal ($if\ n_1 > 10\ \&\ n_2 > 10$) with

- Mean: $\mu_W = \dfrac{n_1(n_1 + n_2 + 1)}{2}$
- Variance: $\sigma_W^2 = \dfrac{n_1 n_2}{12}(n_1 + n_2 + 1)$
- Test Statistic: $z_s = \dfrac{W - \mu_W}{\sigma_W}$

W should be S in the 3 formulas to the left

Rejection Region at the α level of significance:

$H_a: \tilde{\mu}_1 > \tilde{\mu}_2 \quad if \quad z_s \geq z_\alpha$

$H_a: \tilde{\mu}_1 < \tilde{\mu}_2 \quad if \quad z_s \leq -z_\alpha$

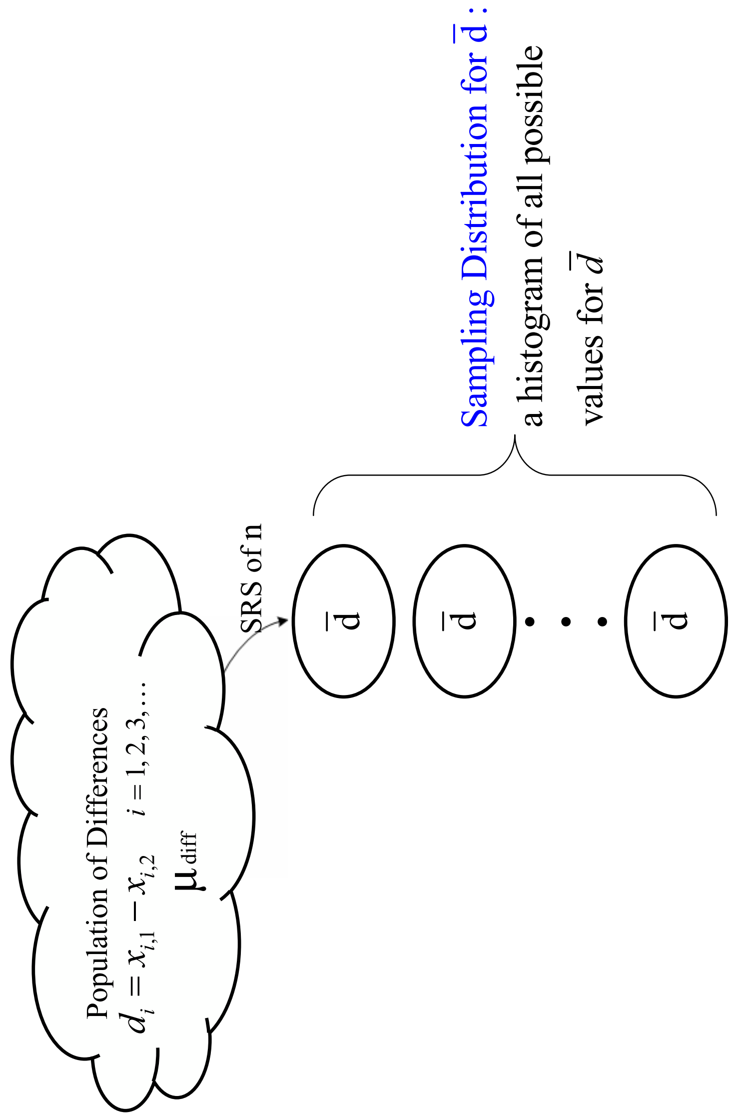
$H_a: \tilde{\mu}_1 \neq \tilde{\mu}_2 \quad if \quad |z_s| \geq z_{\alpha/2}$

When there are ties in the ranks there is a correction to the variance:

$$\sigma_W^2 = \frac{n_1 n_2}{12}(n_1 + n_2 + 1) - \left(\frac{n_1 n_2}{12} \frac{\sum_{j=1}^{k} t_j (t_j^2 - 1)}{(n_1 + n_2)(n_1 + n_2 - 1)} \right)$$

R:

- plac = c(0.9, 0.37, 1.63, 0.83, 0.95, 0.78, 0.86, 0.61, 0.38, 1.97)
  alco = c(1.46, 1.45, 1.76, 1.44, 1.11, 3.07, 0.98, 1.27, 2.56, 1.32)
  wilcox.test(plac, alco, paired=FALSE, correct=F, exact=F)
- When there are more than 2 samples, this is known as the Kruskal-Wallis Test, kruskal.test() in R



T-Test for 2 *Paired* Samples

- Each individual in sample 1 is matched to one individual in sample2 in order to control for other factors and reduce variability
  - $d_i = x_{i,1} - x_{i,2}$ $i = 1, 2, 3, \ldots, n_d$
- $H_o : \mu_d = 0$ → $t_s = \dfrac{\bar{d} - 0}{s_d / \sqrt{n_d}}$ with $df = n_d - 1$
- $100(1-\alpha)\%$ Confidence Interval for μ_d

$$\bar{d} \pm t_{\frac{\alpha}{2}, n_d - 1}\left(\frac{s_d}{\sqrt{n_d}}\right)$$

Example:

Research question: Does oral contraceptive (OC) use increase SBP?

X = Systolic Blood Pressure (SBP)

$d_i = x_{i,(no-OC)} - x_{i,(yes-OC)}$

$H_o : \mu_d = 0$

SBP_no_OC	SBP_yes_OC	diff
115	128	-13
112	115	-3
107	106	1
119	128	-9
115	122	-7
138	145	-7
126	132	-6
109	109	0
104	102	2
115	117	-2

Non-parametric Tests for a single Sample

- Do not rely on the assumption of normally distributed populations → Nonparametric

Wilcoxon Signed Rank Test

- Compute the observe differences, $d_i = x_{1i} - x_{2i}$, for the *n* observed pairs.
- Remove all values with $d_i = 0$
- let n^* be the number of non-zero values.
- Rank the *absolute values* of the differences (ties get assigned the average rank).
  V_+ = Sum of the positive ranks
  V_- = Sum of the negative ranks
- V = minimum (V_+ , V_-)

When H_o is true, the sampling distribution for V is approx. normal (*if* $n > 30$) with

- Mean: $\mu_V = \dfrac{n^*(n^*+1)}{4}$
- Variance: $\sigma_V^2 = \dfrac{n^*(n^*+1)(2n^*+1)}{24}$
- Test Statistic: $z_s = \dfrac{V - \mu_V}{\sigma_V}$

Rejection Region at the α level of significance:
$H_a : \tilde{\mu}_d > 0$ *if* $z_s \geq z_\alpha$
$H_a : \tilde{\mu}_d < 0$ *if* $z_s \leq -z_\alpha$
$H_a : \tilde{\mu}_d \neq 0$ *if* $|z_s| \geq z_{\alpha/2}$

When there are ties in the ranks there is a correction to the variance:

$$\sigma_V^2 = \frac{n^*(n^*+1)(2n^*+1)}{24} - \frac{1}{2}\sum_{j=1}^{k} t_j(t_j - 1)(t_j + 1)$$

R:

- sbp1 = c(115, 112, 107, 119, 115, 138, 126, 109, 104, 115)
  sbp2 = c(128, 115, 106, 128, 122, 145, 132, 109, 102, 117)
  diff = sbp1 - sbp2
- wilcox.test(diff, paired=TRUE, correct=F, exact=F) #OR#
  wilcox.test(sbp1, sbp2, paired=TRUE, paired=TRUE, correct=F, exact=F)

Approximate Sample Size Calculations for 2-Sample Tests (assuming $n_1 = n_2$)

1) For a desired Margin of Error (E):

a) *2 Independent Samples*: $(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}$

$$E \approx z_{\alpha/2}\sqrt{\frac{\hat{\sigma}_1^2}{n} + \frac{\hat{\sigma}_2^2}{n}} \quad \rightarrow \quad n \approx \frac{(z_{\alpha/2})^2\left(\hat{\sigma}_1^2 + \hat{\sigma}_2^2\right)}{E^2}$$

b) *Paired Samples*: $\bar{d} \pm t_{\alpha/2}\frac{\hat{\sigma}_d}{\sqrt{n_d}}$

$$E \approx z_{\alpha/2}\frac{\hat{\sigma}_d}{\sqrt{n_d}} \quad \rightarrow \quad n_d \approx \frac{(z_{\alpha/2})^2\hat{\sigma}_d}{E^2}$$

2) For testing H_o with a fixed Type I and Type II error probability $(\alpha \text{ and } \beta)$

a) *2 Independent Samples*: $vs.\ H_a: \mu_1 - \mu_2 = \Delta$

$$n \approx \frac{\left(\hat{\sigma}_1^2 + \hat{\sigma}_2^2\right)\left(z_\alpha + z_\beta\right)^2}{\Delta^2}$$

($z_{\alpha/2}$ for a 2-sided test)

b) *Paired Samples*: $vs.\ H_a: \mu_d = \Delta$

$$n \approx \frac{\hat{\sigma}_d^2\left(z_\alpha + z_\beta\right)^2}{\Delta^2}$$

($z_{\alpha/2}$ for a 2-sided test)

Ex: (Mouse diets)

Find the number of observations that would be needed to have *power* = 90% for rejecting H_0 in favor of $H_a: \mu_{Bean} - \mu_{Oat} = 7$ at the α=0.05 level (1-sided test).

Descriptive Statistics: Bean, Oat

Variable	N	Mean	StDev	Variance
Bean	15	26.46	5.90	34.81
Oat	15	32.23	9.56	91.45

In a *prospective study* individuals are followed for a period of time to see if a "condition" develops.

- a.k.a. *Longitudinal study* or follow-up study

In a *retrospective study* individuals with & without a "condition" are asked about past exposures (or current status).

- a.k.a.*Cross-sectional study*

Example:

A study of the relationship between oral contraceptive use (OC) and blood pressure (BP) in premenopausal women.

- *Longitudinal study*
  - Identify a group of nonpregnant, women of childbearing age & not current OC users.
  - Measure BP (*baseline measurement*).
  - Re-screen the women after 1 year to determine a subgroup that have become OC users. This subgroup is the *study group*.
  - Measure BP at the follow-up visit
  - Compare the *baseline* and *follow-up* BP of the women in the study group.

- *Cross-sectional study*
  - Identify a group of OC users and a group of non-OC users among nonpregnant, women of childbearing age.
  - Measure the BP of each woman.
  - Compare the BP of the OC and non-OC users.