2. NONPARAMETRIC INFERENCE

The nonparametric approach is based on the fact that we don't want to make "strong" assumptions on the distribution of the data. However, this may imply heavy computational methods.

"strong" assumptions = (e.g.) gaussian data (which are not so frequent in reality)

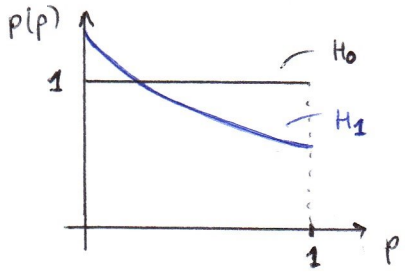
Let's see for instance what happens if we try to use the t-test on non gaussian data (Example 1.). The dataset is small and nongaussian. (⇒ we cannot use t-test and we either can rely on central limit theorem / asymptotic results).

We sample, with sample size $n=5$, from different distributions (normal, uniform and t-student). We generate 100.000 datasets and we apply 100.000 t-tests (and we collect the p-values).

We consider the t-test for one population; assumptions: iid normal data

we want to violate this

t-test: exact test, i.e., under H_0 the distribution of the p-value (which is a random variable depending on the random sample) is the uniform distr. $\mathcal{U}([0,1])$ ⟹ under H_0 all the values of the p-values have the same probability to happen. Under H_1 the distribution of the p-value change:
(if we're not under H_0 we're more likely to see smaller values of p-values)



— depends on the violation of H_0 (in this case it depends on the mean): the longer the difference between the true mean and the tested value of the mean and the more concentred this distribution will be around 0.

idea of the t-test:

we compare the test statistic, we measure the statistical distance between the hypotized mean and our estimate of the mean:

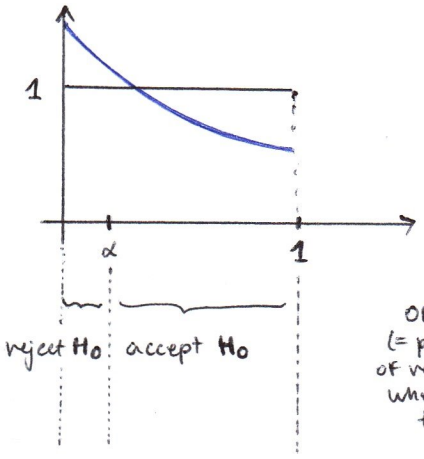
$$= \frac{\text{sample mean} - \text{mean under } H_0}{\text{standard deviation of the estimator}} \sim t(n-1)$$

The p-value is just a measure of the extremity of our sample w.r.t. the normal distribution of the sample under the null hypothesis

small p-value → reject H_0
→ the data are kind of "extreme" w.r.t. the distribution under H_0

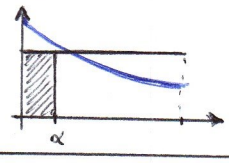
The p-value can be used to run the test: we select a level of significance α and we reject if the p-value is smaller than α.

So if we fix $\alpha = \mathbb{P}(\text{type I error}) = \mathbb{P}(\text{rejecting } H_0 \mid H_0 \text{ is true})$

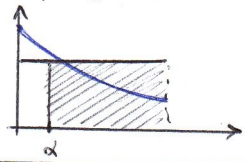


POWER OF THE TEST (= probability of rejecting H_0 when H_0 is false)

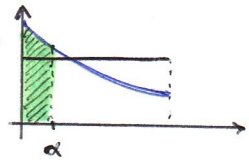
$\mathbb{P}(\text{reject } H_0 \mid H_0 \text{ true}) = \alpha$



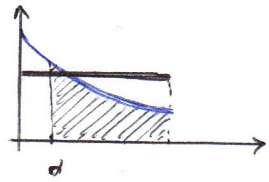
$\mathbb{P}(\text{accepting } H_0 \mid H_0 \text{ T}) = 1-\alpha$



$\mathbb{P}(\text{reject } H_0 \mid H_0 \text{ F}) > \alpha$



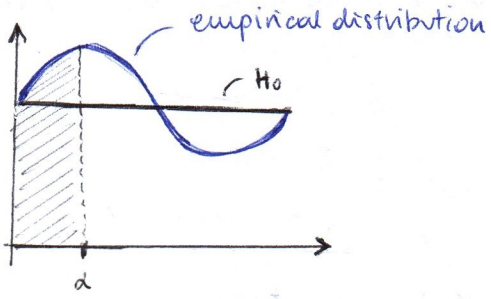
$\mathbb{P}(\text{accept } H_0 \mid H_0 \text{ F}) < 1-\alpha$



So what are we gonna do with the experiment (Example 1) is:
we generate some data under H_0 (= case in which the true mean is identical to the hypothized value) and we estimate the actual probability of rejecting H_0.
So, for each value of α (:= "nominal level") we estimate the empirical level of the test. We'll see if the test is actually rejecting H_0 with probability α or not.

To make this experiment we have to estimate the distribution of the p-value. (We'll do that in Monte Carlo setting). We'll have an estimate of this true distribution of the p-values under H_0 and to compute the empirical level of the test, for a given value α, we'll just have to compute how many p-values are smaller than α.

For instance:



we know that under H_0 ▨ should be α, if it's much higher / smaller then maybe H_0 is not true!
We can check the empirical α and the H_0 α for all the values of $\alpha \in [0,1]$.
(H_0 holds $\Longrightarrow$ ▨ $\sim \alpha$)

In Example 1 we fix $\alpha = 0.05$. Since p-values are uniformly distributed under H_0, having $\alpha = 0.05$ means that, on average 1 time over 20 that we do the test under H_0 true we wrongly reject H_0.
(for instance $\alpha = 0.01$ means 1 time over 100 times that we do the test we wrongly reject H_0).

What can we do in practice if we don't have data that satisfies the assumptions of parametric tests?

- we can try to remove outliers
- we can try to transform data (box-cox transformations)
- non parametric way (↓)

NON-PARAMETRIC STATISTICS

A SIMULATED EXAMPLE ABOUT THE POSSIBLE FAILURE OF PARAM TESTS. (T-TEST FAILURE.R)

SIGN TESTS

(not very used but important since all the other tests are variations of this idea)

(ARBUTHNOT 1710)

we can do this test with an explicit formula for the calculation of the p-value

ONE-SAMPLE (TWO-SIDED) SIGN TEST FOR THE MEDIAN

It simply counts the number of deviations

ASS. $X_1, X_2, \dots, X_m \sim iid(X)$ — generic common distribution (FOR SIMPLICITY CONTINUOUS)

$H_0: MED(X) = c_0 \quad H_1: MED(X) \neq c_0$ ← the test is not based on the mean, it's based on the **MEDIAN** value of X (= value which divides the real axes in two sections of equal probability (50%, 50%))

equivalent ↓

$H_0: P(X > c_0) = 0.5 \quad H_1: P(X > c_0) \neq 0.5$

Test statistic:

$$W = \sum_{i=1}^{m} \mathbb{1}_{\{X_i > c_0\}} \rightsquigarrow \text{NUMBER OF DATA GREATER THAN } c_0$$

$W \in \{0, 1, \dots, m\}$

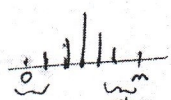
$W \sim B(m, P(X > c_0))$

$W \overset{H_0}{\sim} B(m, 0.5) \qquad W \overset{H_1}{\sim} B(m, p) \text{ WITH } p \neq 0.5$

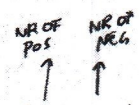
Example: is a male birth more probable than a female one?
X_i = #(male births) − #(female births) during the year i:
$H_0: \mathbb{P}(X_i > 0) = \frac{1}{2}$
$H_1: \mathbb{P}(X_i > 0) \neq \frac{1}{2}$
→ = "they're equiprobable"

If we want to run: $H_0: \mathbb{P}(X > c_0) > 1/2$ then we rely on the right tail, if the test is left sided we use only the left tail of the distribution

THE CRITICAL REGION IS DEFINED BY THE TWO $\alpha/2$-TAIL OF THE $B(m, 0.5)$



THE P-VALUE IS TRIVIALLY DEFINED ACCORDINGLY

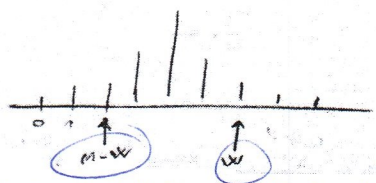


⇒ IF $W^* = \max(W, m - W)$

THE P-VALUE IS THE PROBABILITY OF HAVING MORE (OR EQUAL) THAN W^* HEADS OR MORE (OR EQUAL) THAN W^* TAILS IN m INDEPENDENT TOSSES OF A FAIR COIN.

$$p = 2 \cdot \sum_{k=W^*}^{m} \binom{m}{k} 0.5^m$$

(taken from binomial distribution)



# of values that are less than c_0 — # of values that are greater than c_0

$n - W \overset{H_0}{\sim} Bi(n, \frac{1}{2})$
$W \overset{H_0}{\sim} Bi(n, \frac{1}{2})$

under H_0 both W and n−W are identically distributed ⇒ the inference made on the two is the same

The good thing of this test is that we don't have any assumption on the distribution of the data. The test is valid (and exact) for $\forall$ distribution of data.
However there are some common problems in the nonparametric case:

and so we have some issues with the computation of the quantiles, since it's not always possible to split data in proportions that we want (2 data? We can't have 1/3, 2/3)

(NB) THE DISTRIBUTION OF W IS **DISCRETE**

$\Rightarrow$ THE SAMPLE SIZE m DETERMINES THE SO-CALLED **ACHIEVABLE SIGN LEVELS OF THE TEST**

(i.e. LEVELS OF SIGNIFICANCE FOR WHICH AN α-SIZE TEST CAN BE BUILT WITHOUT A RANDOMIZATION STRATEGY)

If we select 10% we would like to have a tail of 5% and an other tail of 5%. However we cannot do that. We can take a tail of 0%, or 15.6%, or 31.2%, or 3,1%, not others.
$\Rightarrow$ Only 4 possible values.
These values are the ACHIEVABLE LEVELS of the test (/reachable levels) = values for which we can give an exact test ($\mathbb{P}$(reject H_0) = desired value)

ex $m=5$

$P(W=0) = 3.1\%$
$P(W=1) = 15.6\%$
$P(W=2) = 31.2\%$
$P(W=3) = 31.2\%$
$P(W=4) = 15.6\%$
$P(W=5) = 3.1\%$

ACHIEVABLE LEVELS OF α ARE
$\alpha = 0\%$
$\alpha = 6.2\% = 2 \cdot 3.1\%$ $\longrightarrow$ reject if $W=0$ or $W=5$
$\alpha = 37.4\% = 2 \cdot (3.1\% + 15.6\%)$ $\longrightarrow$ reject if $W=0,5,1,4$
$\alpha = 100\%$

IN ALL OTHER CASES α
- α CAN BE FLOORED TO LOWER AND CLOSEST ACHIEVABLE VALUE (OBTAINING A CONSERVATIVE TEST) *
- A RANDOMIZATION STRATEGY CAN BE PUT ON TOP OF THE PROCEDURE **

(NB) IF THE DISTRIBUTION OF X IS SIMMETRIC AND THE MEAN EXISTS $\Rightarrow$ MED(X) $\equiv$ MEAN(X)

In this case the sign test is also a test for the **mean**.

THIS IS GENERALLY NOT TRUE
ex. $X \sim Exp(\lambda)$
$MEAN(X) = 1/\lambda$
$MEDIAN(X) = \log 2/\lambda = \frac{0.693}{\lambda}$

(NB) FOR LARGE VALUE OF m: $B(m, 0.5) \approx N(m/2, m/4)$

(NB) W CAN BE ALSO COMPUTED AS $W = \frac{m + \sum_{i=1}^{m} sign(X_i - C_0)}{2}$

(NB) ONE-SIDED TESTS CAN BE TRIVIALLY BUILT FOCUSING ON THE CORRESPONDING TAIL.

ex RIGHT-SIDED
$H_0: MED(X) = C_0 \quad H_1: MED(X) > C_0$
$[\, H_0: P(X > C_0) = 0.5 \quad H_1: P(X > C_0) > 0.5 \,]$

$W = \sum_{i=1}^{m} \mathbb{1}_{\{X_i > C_0\}} \qquad P = \sum_{k=w}^{m} \binom{m}{k} 0.5^m$

ACHIEVABLE LEVELS ARE DIFFERENT FROM THE ONE-SIDED CASE

(NB) EXTENSIONS TO THE CASE OF DISCRETE OR ORDINAL DATA ARE POSSIBLE. (REMOVAL OF $X_i = C_0$)

* $\alpha = 10\%$ $\longrightarrow$ left tail and right tail 5%, we take 3,1% $\longrightarrow$ we obtain a value of 6.2%, which is not exact but $\mathbb{P}$(reject H_0 | H_0 true) < 10% (even if we would like to have $\alpha = \alpha$)

** Used when we want to make continuous a discrete variable.
If we want to obtain $\alpha = 10\%$ we consider:
- reject if $W=0$ (3.1%)
- reject if $W=5$ (3.1%)
- if $W=1$ or $W=4$ we toss a coin with a given probability s.t. the general hypothesis lead us to a global probability of rejection of 10%

TWO-SAMPLE PAIRED SIGN TESTS

ASS. $X_1-Y_1, X_2-Y_2, \dots, X_m-Y_m$ ~ ind $Z = X-Y$ (CONTINUOUS)

$H_0: MED(X-Y) = 0 \quad H_1: MED(X-Y) \neq 0$

$[H_0: P(X>Y) = 0.5 \quad H_1: P(X>Y) \neq 0.5]$

Test statistic:

$$W = \sum_{i=1}^{m} \mathbb{1}_{\{X_i > Y_i\}} = \frac{m + \sum_{i=1}^{m} \text{sign}(X_i - Y_i)}{2} \quad \text{ecc.}$$

(N.B.) THE TEST CAN BE ALSO NOT CENTERED ON "0"

(N.B.) WE ARE NOT TESTING THE EQUALITY IN DISTRIBUTION OF X AND Y BUT THE HYPOTHESIS THAT IN A RANDOM PAIR X HAS THE SAME PROBABILITY OF BEING LARGER OR SMALLER THAN Y.
(OF COURSE WITH OTHER ASS, for example symmetry, so we could end up testing mean, or other;...)

GENERAL COMMENTS

- SIGN TESTS DO NOT REQUIRE NORMALITY → because it's not like the t-test, we're not comparing means!
- " " ARE ROBUST TO:
  - VIOLATION OF NORMALITY (≠ t-test)
  - PRESENCE OF OUTLIERS
- SIGN TEST ARE LESS POWERFUL THAN THE "OPTIMAL" PARAMETRIC TEST

the presence of outliers influences the mean, but here we're not dealing with the mean: an outlier simply means that $W \to W+1$

- R EXAMPLE SIGN-TEST.R (Example 2)

Weakness: the fact that $X = 1 > 0$ or $X = 1.000.000 > 0$ doesn't change at all the W statistics, which in all cases take a value of 1 (which however helps with the robustness with outliers).

This test's power is also a weakness: this test is not able to distinguish a major deviation from C_0 from a minor one.

This weakness pushed scientists to work out a solution which takes into account the magnitude of deviation, not only the sign.

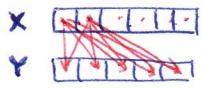
They looked for something:

- distribution independent
- taking into account the magnitude of deviation

→ this lead to:

1. MANN-WHITNEY TEST
2. WILCOXON TEST

* The way to take into account the magnitude of the deviation without introducing assumptions on the distributions is thruog "contests". We create a contest between the X team and the Y team. We have n_1 players for the first team and n_2 players for the second. We build all the possible contests: (how many contests? $n_1 \cdot n_2$)



$\Rightarrow$ U_1 = # contests won by X $\Rightarrow$ $\frac{U_1}{n_1 \cdot n_2}$ = max likelihood estimator of $\mathbb{P}(X>Y)$

U_2 = # contests won by Y $\Rightarrow$ $\frac{U_2}{n_1 \cdot n_2}$ = max likelihood est. of $\mathbb{P}(Y>X)$

RANK TESTS

non-parametric counterpart for the t-test for two populations

MANN-WHITNEY U-TEST (TWO-SAMPLE RANK SUM TEST) (1945 WILCOXON) (1947 MANN & WHITNEY)

ASS. $X_1, X_2, \dots, X_{m_1} \sim$ ind X (CONTINUOUS)
$Y_1, Y_2, \dots, Y_{m_2} \sim$ ind Y (CONTINUOUS) $\Big\rangle$ INDEPENDENT

$H_0: P(X>Y) = 0.5 \quad H_1: P(X>Y) \neq 0.5$ probability that the first random variable (X) is larger than the second (Y)

Test statistics:

$$U_1 = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \mathbb{1}_{\{X_i > Y_j\}} \equiv m_1 m_2 \hat{P}(X>Y)$$ [NUMBER OF PAIRWISE "CONTESTS" WON BY THE FIRST SAMPLE]

$$U_2 = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \mathbb{1}_{\{Y_j > X_i\}} \equiv m_1 m_2 \hat{P}(Y>X)$$ [NUMBER OF PAIRWISE "CONTESTS" WON BY THE SECOND SAMPLE]

(NB) $U_1 + U_2 = \underbrace{m_1 m_2}_{\text{NR OF CONTESTS}}$ $\Big\rangle$ THE INFO PROVIDED BY THE TWO TEST STATISTICS IS THE SAME! $\Rightarrow$ we can choose one or the other to make inference

U_1 AND U_2 CAN BE COMPUTED IN A MORE EFFICIENT WAY AS FOLLOWS:

However sometimes it gets computationally heavy, so we procede in another way to find U_1 and U_2:

- DEFINE $\mathbb{X} = (X_1, X_2, \dots, X_{m_1}, Y_1, Y_2, \dots, Y_{m_2})$, $\dim(\mathbb{X}) = n_1 + n_2$
  (POOLED SAMPLE) = concatenation of the samples from X and Y
- COMPUTE THE RANK OF THE m_1+m_2 DATA WRT THE POOLED SAMPLE

If a datum is ranked #1 then it's the worst in the game (opposite than intuitively)

$$\mathbb{R} = (\pi(X_1), \dots, \pi(X_{m_1}), \pi(Y_1), \dots, \pi(Y_{m_2}))$$

$$\pi(X_i) = \sum_{x \in \mathbb{X}} \mathbb{1}_{\{x \leq X_i\}} \quad i = 1, \dots, m_1$$ [NR OF WINS OR TIES OF X_i AGAINST $\mathbb{X}$]

$$\pi(Y_j) = \sum_{x \in \mathbb{X}} \mathbb{1}_{\{x \leq Y_j\}} \quad j = 1, \dots, m_2$$ [NR OF WINS OR TIES OF Y_j AGAINS $\mathbb{X}$]

- COMPUTE THE SUM OF THE RANKS OF THE TWO SAMPLES.

(**)
$$R_1 = \sum_{i=1}^{m_1} \pi(X_i) = \sum \text{ranks of the 1st sample}$$
$$R_2 = \sum_{i=1}^{m_2} \pi(Y_j) = \sum \text{ranks of the 2nd sample}$$

- $$U_1 = R_1 - \frac{m_1(m_1+1)}{2} \qquad U_2 = R_2 - \frac{m_2(m_2+1)}{2}$$

these formulas are the ones that give the name "rank sum test"

R_1: # wins/ties of the first sample against the pooled sample

$\frac{m_1(m_1+1)}{2}$: # wins/ties of the first sample against itself

$\Rightarrow$ U_i $(i=1,2)$ is simply computed by adjusting R_i

(**): $\frac{(n_1+n_2)(n_1+n_2+1)}{2} = R_1 + R_2 = \sum$ of all the ranks

Note: Under H_1 $\mathbb{X}$ is not a "proper" sample since n_1 elements are from 1 distribution (X) and n_2 are from the 2nd (Y). Under H_0 instead it is drawn from the same distribution.

H_0 true ⟹ The distribution of X and Y are the same. In this case data are iid and so each datum has the same probability to be ranked in $\forall$ position (this happens for all the distributions X and Y)

⟹ under H_0 the distribution of U_1 and U_2 do not depend on the distribution of the data!

WE DO NOT HAVE AN EXPLICIT FORMULATION OF THE DISTRIBUTION OF U_1 (OR U_2) UNDER H_0!

HOWEVER THE DISTRIBUTION IS ALWAYS THE SAME:

- JUST DEPENDING ON m_1 AND m_2
- NOT DEPENDING ON THE (SAME) DISTRIBUTION OF X AND Y (PROOF: $\pi(x_i) \overset{\#}{\sim} U\{1, \dots, m_1+m_2\}$)

⇒ QUANTILE CAN EASILY ESTIMATED VIA A ONE-TIME MC SIMULATION

MOREOVER: UNDER H_0:

- U_1 AND U_2 HAS THE SAME (DISCRETE) DISTRIBUTION $\left(U_1 - \frac{m_1 m_2}{2} = -\left(U_2 - \frac{m_1 m_2}{2}\right)\right)$
- U_1 AND U_2 ARE SYMMETRICALLY DISTRIBUTED AROUND $\frac{m_1 m_2}{2}$
  - $E(U_1) = E(U_2) = \frac{m_1 m_2}{2}$
  - $VAR(U_1) = VAR(U_2) = \frac{m_1 m_2 (m_1 + m_2 + 1)}{12}$

U_1 and U_2 depends only on n_1 and n_2 (under H_0)

⇒ CRITICAL REGIONS CAN BE COMPUTED ACCORDINGLY.

⇒ P-VALUE " " " "

RELYING ON THE QUANTITY $U^* = \max(U_1, U_2) =$

$= \max(U_1, m_1 m_2 - U_1) =$

$= \max(U_2, m_1 m_2 - U_2)$

(NB) BEING U_1 (AND U_2) $\overset{\text{UNDER } H_0}{\vee}$ THE SUM OF EXCHANGEABLE RANDOM VARIABLES (i.e. THE RANKS) WITH FINITE MEAN AND VARIANCE AND VANISHING COVARIANCES THE CLT HOLDS:

$$\frac{U_1 - \frac{m_1 m_2}{2}}{\sqrt{\frac{m_1 m_2 (m_1 + m_2 + 1)}{12}}} \overset{m_1, m_2 \to +\infty}{\sim} N(0,1)$$

⟩ NORMAL APPROXIMATIONS OF THE QUANTILES.

(NB) ONE-SIDED TESTS CAN BE TRIVIALLY BUILT RELYING ON THE CORRESPONDING TAIL.

(NB) THE U-TEST CAN BE ADAPTED ALSO TO DISCRETE AND ORDINAL DATA (i.e. AVERAGE RANK)

(NB) THE U-TEST IS INVARIANT WRT MONOTONIC TRANSFORMATION OF DATA. (PROOF: RANKS ARE INVARIANT) NO NEED FOR Box. Cox TRANSFORM

R EXAMPLE MANN-WHITNEY U TEST R

proposal that tries to solve the issue of "weightening" different magnitude of deviations from c_0 (the idea starts from the sign test)

WILCOXON SIGNED-RANK TEST (WILCOXON 1945)

ONE-SAMPLE TWO-SIDED TEST (CENTERED ON "0").

IT IS AN EXTENSION OF THE SIGN-TEST IN WHICH ALSO THE MAGNITUDE (i.e. RANK OF THE ABSOLUTE VALUE) IS TAKEN INTO CONSIDERATION

ASS. $X_1, X_2, \dots, X_n \sim_{iid} X$ (CONTINUOUS)

$H_0: P(X>0) = 0.5 \quad H_1: P(X>0) \neq 0.5$ ← same test of the sign test (we're stating "median = 0")

what changes is the test statistics:

$$W^+ = \sum_{i=1}^{n} \mathbb{1}_{\{X_i>0\}} R(|X_i|)$$

→ SUM OF THE "ABSOLUTE" RANKS OF POSITIVE VALUES...

$R(|X_i|)$: RANK OF THE ABSOLUTE VALUES

→ the weights depend on the magnitude of deviation of data

- UNDER THE H_0 (INDEPENDENTLY ON THE DISTRIBUTION OF X)
  - $\mathbb{1}_{\{X_i>0\}} \overset{IND}{\sim} B(1, 0.5)$
  - $R(|X_i|) \overset{EXC}{\sim}{}^{*} U(\{1, 2, \dots, n\})$

  ⟩ INDEPENDENT

  The only dependence from $X_1, \dots, X_n \sim X$ is n. That's why W^+ depends only on n.

  SO THE DISTRIBUTION OF W^+ UNDER H_0 DEPENDS ONLY ON n

  ⇒ QUANTILES CAN BE ESTIMATED VIA A ONE-TIME MC SIMULATION (EXPLICIT FORMULA IS NOT AVAILABLE FOR GENERIC n)

* exchangeable (≠ iid): a given vector of a random variable is exchangeable if after permutation of components ($\forall$ permutation) we obtain a random vector with exactly the same joint distribution

iid ⇄ exchang.

- UNDER H_1:
  - $\mathbb{1}_{\{X_i>0\}} \overset{IND}{\sim} B(1, p)$ WITH $p \neq 0.5$
  - $R(|X_i|) \overset{EXC}{\sim} U(\{1, 2, \dots, n\})$

  ⟩ DEPENDENT

If we have that the distribution is around zero then any datum can be ranked anywhere with the same probability. If the distribution has positive median we have that values associated to larger ranking (= associated to larger deviation) are more likely to be positive than negative.

Another way to do the same test:

SIMILARLY AND EQUIVALENTLY:

$$W^- = \sum_{i=1}^{n} \mathbb{1}_{\{X_i<0\}} R(|X_i|)$$

→ SUM OF THE "ABSOLUTE" RANKS OF THE NEGATIVE VALUES.

$W^+ + W^- = \frac{n(n+1)}{2}$ = summation of the ranks of all data

- W^+ AND W^- HAS THE SAME DISTRIBUTION SYMMETRIC AROUND $\frac{n(n+1)}{4}$

SO INFERENCE CAN BE EQUIVALENTLY CARRIED OUT RELYING ON:

- W^+
- W^-

  ⟩ CENTERED ON $\frac{n(n+1)}{4}$ ⟩ P-VALUE BASED ON $\max(W^+, W^-)$

- $W^+ - W^- = \frac{n(n+1)}{2} - 2W^- = \sum_{i=1}^{n} \text{sign}(X_i) R(|X_i|)$
- $W^- - W^+ = \frac{n(n+1)}{2} - 2W^+ = \sum_{i=1}^{n} \text{sign}(X_i) R(|X_i|)$

  ⟩ CENTERED ON "0" ⟩ P-VALUE BASED ON $|W^+ - W^-|$

that's why they're called "sign rank test"

- THE NON ZERO-CENTERED VERSION CAN BE EASILY IMPLEMENTED REPLACING $X_i \mapsto X_i - c_0$

general test (we can also change 0.5)

$H_0: P(X > c_0) = 0.5 \quad H_1: P(X > c_0) \neq 0.5$

$$W^+ = \sum_{i=1}^{m} \mathbb{1}_{\{X_i > c_0\}} R(|X_i - c_0|)$$

- THE ONE-SIDED VERSION CAN BE BUILT RELYING ON THE CORRESPONDING TAIL.

(VIB) THE SIGNED-RANK TEST IS USUALLY
- MORE POWERFUL THAN THE SIGN-TEST
- LESS " " " OPTIMAL PARAMETRIC TEST

(VIB) IT IS LESS ROBUST THAN THE SIGN-TEST. (w.r.t. outliers)
BUT MORE ROBUST THAN THE OPTIMAL PARAMETRIC TEST
(REBALANCING OF THE ABSOLUTE RANKS)

TWO-SAMPLE (PAIRED) SIGNED-RANK TEST (CENTERED ON 0)

ASS. $X_1 - Y_1, X_2 - Y_2, \dots, X_m - Y_m \sim \text{ind } Z = X - Y$ (CONTINUOUS)

$H_0: P(X > Y) = 0.5 \quad H_1: P(X > Y) \neq 0.5$

$[\text{MED}(X-Y) = 0] \quad [\text{MED}(X-Y) \neq 0]$

$$W^+ = \sum_{i=1}^{m} \mathbb{1}_{\{X_i > Y_i\}} R(|X_i - Y_i|)$$

→ SUM OF THE ABSOLUTE RANKS OF THE PAIRS IN WHICH $X_i > Y_i$

R EXAMPLE WILCOXON SIGNED-RANK TEST. R

Conditions:
we know the distributions under H_0 and H_1. (moreover an optimal test can be identified)

Ranking with discrete data?
If X_i is ranked 3 and x_j is ranked 4
$\Rightarrow$ x_i, x_j both ranked 3.5

Sign rank test with discrete data?
if we have a 0 (an observation which is 0) ~~cannot~~ we cannot assign a sign $\Rightarrow$ we simply remove the data