ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ

ΚΕΦΑΛΑΙΟ 3: ΑΝΑΛΥΣΗ ΔΥΟ ΜΕΤΑΒΛΗΤΩΝ

ΣΤΑΤΙΣΤΙΚΗ ΣΥΜΠΕΡΑΣΜΑΤΟΛΟΓΙΑ ΣΤΗ ΣΥΓΚΡΙΣΗ ΔΥΟ ΔΕΙΓΜΑΤΩΝ

ΕΙΣΑΓΩΓΗ: Έστω ότι θέλουμε να συγκρίνουμε δύο πληθυσμούς δηλ. δύο σύνολα ατόμων, φαινομένων, καταστάσεων κλπ, γενικότερα υποκειμένων στα οποία παρατηρούμε κάποιο χαρακτηριστικό που μας ενδιαφέρει. Η σύγκριση των πληθυσμών έχει στόχο τη διατύπωση ύπαρξης ή όχι διαφορών μεταξύ τους με τρόπο τεκμηριωμένο λαμβάνοντας υπόψη τα σφάλματα τυχαιότητας που υπεισέρχονται στις μετρήσεις. Π.χ. θέλουμε να διαπιστώσουμε την ύπαρξη ή όχι διαφορών σε δύο μεθόδους διδασκαλίας ή σε δύο διαφορετικές προπονητικές τεχνικές. Από κάθε πληθυσμό λαμβάνουμε ένα τυχαίο δείγμα και έτσι έχουμε δύο ομάδες παρατηρήσεων. Π.χ. στους μαθητές της ίδιας τάξης εφαρμόζονται δύο διαφορετικές τεχνικές διδασκαλίας για ένα συγκεκριμένο μάθημα. Π.χ. παρατηρήσεις είναι οι βαθμοί των μαθητών στο test που δόθηκε στο τέλος των παραδόσεων. Οι μαθητές του κάθε σχολείου σχηματίζουν τα δείγματα από τους αντίστοιχους υποθετικούς πληθυσμούς.

Οι στατιστικές τεχνικές για τη σύγκριση δύο πληθυσμών, χωρίζονται σε δύο μεγάλες κατηγορίες, τις παραμετρικές και τις μη-παραμετρικές.

- Οι <u>παραμετρικές</u> απαιτούν την ύπαρξη συγκεκριμένων συνθηκών για τα μεγέθη των δειγμάτων ή του πληθυσμού από τον οποίον προέρχονται. Ο έλεγχος που πραγματοποιούν αφορά παραμέτρους του πληθυσμού (π.χ. μέση τιμή, διακύμανση). Το ενδιαφέρον εστιάζεται π.χ. στη σύγκριση της μέσης τιμής ενός πληθυσμού με τη μέση τιμή ενός άλλου, π.χ. μέσα εισοδήματα κατοίκων δύο περιοχών, μέση επίδοση μαθητών σε δύο σχολεία

Ελέγχουμε: $H_0: \mu_1 - \mu_2 = 0$ vs $H_1: \mu_1 - \mu_2 \neq 0$

- Οι μη-παραμετρικές δεν απαιτούν υποθέσεις για τους πληθυσμούς από τους οποίους προέρχονται τα δείγματα και ούτε έχουν περιορισμούς για τα μεγέθη τους.

Ελέγχουμε: H_0: Τα δύο δείγματα προέρχονται από τον ίδιο πληθυσμό
H_1: τα δύο δείγματα προέρχονται από διαφορετικούς πληθυσμούς

Η επιλογή της στατιστικής τεχνικής εξαρτάται από τον σχεδιασμό και το σκοπό της μελέτης αλλά κυρίως από τη φύση των δεδομένων.

Κρίσιμα ερωτήματα αφορούν:
- την ανεξαρτησία των δειγμάτων
- την κατανομή των δεδομένων: Οι παραμετρικές τεχνικές σε ανεξάρτητα δείγματα απαιτούν την ύπαρξη της κανονικής κατανομής του πληθυσμού από τον οποίο προέρχονται.

Επίσης, αν τα δείγματα είναι $n \geq 30$ (αρκετά μεγάλα) το ΚΟΘ εξασφαλίζει την ικανοποίηση αυτών των προϋποθέσεων.

Όταν τα δείγματα είναι συσχετισμένα η υπόθεση της κανονικότητας ελέγχεται για τη διαφορά των μετρήσεων $X_i - Y_i = d_i$

- τον χαρακτήρα των δεδομένων: Ποιοτικά ή ποσοτικά.

Δεν μπορούμε να εφαρμόσουμε παραμετρικές τεχνικές στην περίπτωση διατεταγμένων ποσοτικών δεδομένων

<u>Ανεξάρτητα δείγματα - Το παραμετρικό t-test</u>

- Διασπορές σ_1^2, σ_2^2 γνωστές ή άγνωστες $n, m \geq 30$

Μονόπλευρο test

$H_0: \mu_1-\mu_2=\delta$
$H_1: \mu_1-\mu_2>\delta$

$R=\{z>z_\alpha\}$

ή

$H_0: \mu_1-\mu_2=\delta$
$H_1: \mu_1-\mu_2<\delta$

$R=\{z<-z_\alpha\}$

Αμφίπλευρο test

$H_0: \mu_1-\mu_2=\delta$
$H_1: \mu_1-\mu_2\neq\delta$

$R=\{|z|>z_{\alpha/2}\}$

Στ. Συνάρτηση ελέγχου

$$z=\frac{\bar{x}-\bar{y}-\delta}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}} \quad \text{ή} \quad z=\frac{\bar{x}-\bar{y}-\delta}{\sqrt{\frac{s_1^2}{n}+\frac{s_2^2}{m}}}$$

- $n,m\leq 30$, $\sigma_1^2=\sigma_2^2=\sigma^2$ (Θεωρούμε ότι οι πληθυσμοί από τους οποίους λαμβάνονται τα δείγματα είναι κανονικοί)

Μονόπλευρο test

$H_0: \mu_1-\mu_2=\delta$
$H_1: \mu_1-\mu_2>\delta$

$R=\{t>t_{n+m-2;\alpha}\}$

ή

$H_0: \mu_1-\mu_2=\delta$
$H_1: \mu_1-\mu_2<\delta$

$R=\{t<-t_{n+m-2;\alpha}\}$

Αμφίπλευρο test

$H_0: \mu_1-\mu_2=\delta$
$H_1: \mu_1-\mu_2\neq\delta$

$R=\{|t|>t_{n+m-2;\alpha/2}\}$

Στ. Συνάρτηση Ελέγχου

$$t=\frac{\bar{x}-\bar{y}-\delta}{\sqrt{\frac{(n-1)s_1^2+(m-1)s_2^2}{n+m-2}}\sqrt{\frac{1}{n}+\frac{1}{m}}}$$

Για να ελεγχθεί η ισότητα των διακυμάνσεων στους δύο πληθυσμούς κάνουμε τους ακόλουθους ελέγχους υποθέσεων

Μονόπλευρο test

$H_0: \frac{\sigma_1^2}{\sigma_2^2}=1$ ή $H_0: \sigma_1^2=\sigma_2^2$

$H_1: \frac{\sigma_1^2}{\sigma_2^2}>1$ ή $H_1: \sigma_1^2>\sigma_2^2$

$R=\{f>f_{\nu_1,\nu_2;\alpha}\}$

$f=\frac{s_1^2}{s_2^2}$

ή

$H_0: \frac{\sigma_1^2}{\sigma_2^2}=1$

$H_1: \frac{\sigma_1^2}{\sigma_2^2}<1$

$R=\{f>f_{\nu_1,\nu_2;\alpha}\}$

$f=\frac{s_2^2}{s_1^2}$

Αμφίπλευρο test

$H_0: \dfrac{\sigma_1^2}{\sigma_2^2} = 1$

$H_1: \dfrac{\sigma_1^2}{\sigma_2^2} \neq 1$

$F = \dfrac{s_1^2}{s_2^2}$, όταν $s_1^2 > s_2^2$

$F = \dfrac{s_2^2}{s_1^2}$, όταν $s_2^2 > s_1^2$

$R = \{F > F_{\nu_1, \nu_2; \alpha/2}\}$

ν_1: β.ε. αριθμητή, ν_2: β.ε. παρονομαστή

Προϋπόθεση: Τα δείγματα προέρχονται από κανονικούς πληθυσμούς.

Παράδειγμα Εταιρεία προστασίας καταναλωτών θέλει να ελέγξει αν η διάρκεια ζωής των ελαστικών αυτοκινήτων είναι ίδια για δύο μάρκες A και B. Ένα τ.δ. 8 προϊόντων από κάθε μάρκα ελέγχθηκε σε προσομοιωμένες οδικές συνθήκες και καταγράφηκε ο αριθμός των χιλιομέτρων πριν να σημειωθεί φθορά στο πέλμα. Τα αποτελέσματα σε χιλιάδες Km ήταν:

A	x_i	40.6	42.8	48.4	45.2	39.7	47.4	49.2	44.5
B	y_i	40.9	45.6	38.4	35.8	47.5	42.6	51.4	50.2

Να ελεγχθεί σε $\alpha = 0.1$ η υπόθεση της ισότητας των διακυμάνσεων στους δύο πληθυσμούς, αν οι αντίστοιχες κατανομές τους μπορεί να υποτεθούν κανονικές. Με βάση αυτό το αποτέλεσμα να εφαρμοστεί το t-test για να ελεγχθεί για $\alpha = 0.1$ η ισότητα των δύο μέσων τιμών πληθυσμού για τις δύο μάρκες A και B.

Λύση Έστω X_1, X_2 η διάρκεια ζωής σε χιλιάδες Km του ελαστικού της A και B μάρκας, αντίστοιχα. Ελέγχουμε:

$H_0: \dfrac{\sigma_1^2}{\sigma_2^2} = 1$ vs $H_1: \dfrac{\sigma_1^2}{\sigma_2^2} \neq 1$

$\bar{x}_1 = \frac{\sum_{i=1}^{8} x_{1i}}{8} = 44.7$, $\bar{x}_2 = \frac{1}{8}\sum_{i=1}^{8} x_{2i} = 44.05$

$$s_1^2 = \frac{\sum_{i=1}^{8} x_{1i}}{7} - \left(\frac{8}{7}\right)\bar{x}_1^2 = 12.42 \Rightarrow \underline{s_1 = 3.52}$$

$$s_2^2 = \frac{\sum_{i=1}^{8} x_{2i}}{7} - \left(\frac{8}{7}\right)\bar{x}_2^2 = 31.17 \Rightarrow s_2 = 5.58$$

$$f = \frac{s_1^2}{s_2^2} = \frac{12.42}{31.17} = 0.398$$

Από τον πίνακα της κατανομής $\mathcal{F}$ έχουμε: $\mathcal{F}_{7,7;0.05} = 3.79 \Rightarrow$

$$\Rightarrow \underset{n_1\, n_2;\, 1-\alpha}{\mathcal{F}_{7,7;0.95}} = \frac{1}{\underset{n_2, n_1;\, \alpha}{\mathcal{F}_{7,7;0.05}}} = 0.26$$

Επειδή $\left.\begin{array}{l} f < \mathcal{F}_{7,7;0.05} \\ \text{ή } f > \mathcal{F}_{7,7;0.95} \end{array}\right\} \Rightarrow$

$\Rightarrow$ η H_0 <u>δεν</u> μπορεί να απορριφθεί σε ε.σ.σ. $\alpha = 0.05$.

Ο έλεγχος για την ισότητα των δύο μέσων τιμών πληθυσμού μπορεί να γίνει με το t-test θεωρώντας ότι οι κατανομές των πληθυσμών είναι κανονικές και ότι: $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

Ελέγχουμε $H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2$

Είναι $t_{n_1+n_2-2;\alpha/2} = t_{14;0.05} = 1.761$. Εκτιμούμε την κοινή διακύμανση του πληθυσμού με τη στ.σ.

$$s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{7(12.42) + 7(31.17)}{14}$$

$$= 21.795 \Rightarrow \underline{s = 4.67}$$

Η στ. συνάρτηση ελέγχου είναι:

$$|t| = \frac{|\bar{x}_1 - \bar{x}_2|}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{44.7 - 44.05}{4.67\sqrt{\frac{1}{8} + \frac{1}{8}}} = 0.28$$

Όμως $t < t_{14;0.05}$ και συνεπώς η H_0 δεν μπορεί να απορριφθεί.

Λύση με το SPSS Αρχικά θα ελέγξουμε την υπόθεση των ίσων διακυμάνσεων και στη συνέχεια θα εφαρμόσουμε το παραμετρικό t-test.

Analyze → Descriptive Statistics → Explore

Plots (Normality plots with tests)

Εστιάζουμε στα tests of Normality: K-S και Shapiro-Wilk Και από τα δύο tests και για τα δύο δείγματα A, B φαίνεται ότι προέρχονται από κανονικούς πληθυσμούς. Ισχυρές "γραφικές" ενδείξεις λαμβάνουμε και από τα γραφήματα Q-Q (Quantile-Quantile) και detrended normal Q-Q. Η απόκλιση των δεδομένων από τη διχοτόμο y-x ($1^{ης}$, $3^{ης}$ γωνίας των αξόνων) που αντιστοιχεί στην κανονική κατανομή είναι "μικρή". Ισχυρή ένδειξη κανονικότητας των πληθυσμών από τον οποίον προέρχονται τα δύο δείγματα.

Έχουμε τις προϋποθέσεις για να προχωρήσουμε στην εφαρμογή του t-test.

Analyze → Compare Means → Independent Samples T-Test

Test Variable (Duration)

Grouping Variable (Markes)

Define Groups Group 1: A
Group 2: B

Στον Πίνακα Group Statistics λαμβάνουμε αρχικά μία μικρή περιγραφική στατιστική των παρατηρήσεων και στον δεύτερο Πίνακα Independent Samples Test βλέπουμε τα αποτελέσματα του στατιστικού ελέγχου T-test.

Ο έλεγχος της ισότητας των διακυμάνσεων γίνεται με το test του Levene (Levene's Test for Equality of Variances).

Levene's test (1960) is used to test if K samples have equal variances. Equal variances across samples is called homogeneity of variance. Some statistical tests, for example, the analysis of variance (ανάλυση διακύμανσης) assume that variances are equal across groups or samples. The Levene test can be used to verify that assumption.

The Levene test is defined as:

$H_0: \sigma_1 = \sigma_2 = \dots = \sigma_K$

$H_1: \sigma_i \neq \sigma_j$ for at least one pair (i,j)

Test Statistic: Given a variable Y with sample of size N divided into K-subgroups, where N_i is the sample size of the i-th subgroup, the Levene test statistic is defined as:

$$W = \frac{(N-K)\sum_{i=1}^{K} N_i(\bar{Z}_{i.} - \bar{Z}_{..})^2}{(K-1)\sum_{i=1}^{K}\sum_{j=1}^{N_i}(Z_{ij} - \bar{Z}_{i.})^2}$$

where Z_{ij} can have one of the following three definitions:

① $Z_{ij} = |Y_{ij} - \overline{Y}_{i.}|$, where $\overline{Y}_{i.}$ is the mean of the i-th subgroup.

② $Z_{ij} = |Y_{ij} - \overline{Y}_{i.}|$, where $\overline{Y}_{i.}$ is the median of the i-th subgroup

③ $Z_{ij} = |Y_{ij} - \overline{Y}'_{i.}|$, where $\overline{Y}'_{i.}$ is the 10% trimmed mean of the i-th subgroup.

$\overline{Z}_{i.}$ is the group means of the Z_{ij} and $Z_{..}$ is the overall mean of the Z_{ij}. The definition of the test based on the median is recommended as the choice that provides good power of the test.

The Levene's test rejects the hypothesis that the variances are equal if $W > F_{K-1, N-K; \alpha}$ where $F_{K-1, N-K, \alpha}$ is the upper critical value of the F-distribution with K-1 and N-K degrees of freedom at a significance level α.

Εδώ: $F = W = 2{,}587$, $\hat{\alpha} = 0{,}130$ (Ισχυρές ενδείξεις μη-απόρριψης της H_0)
(για Levene's test)

$t = 0.289$, $df = 14$, $\hat{\alpha} = 0.777$ (Ισχυρές ενδείξεις μη-απόρριψης της H_0)
(για t-test)

Mean difference (Διαφορά των μέσων τιμών στα δύο δείγματα) $= \bar{x}_2 - \bar{x}_1 = 0{,}6750$

Std. Error difference $= S_2 - S_1 = 5.58 - 3.52 = 2.3341$

95% CI for the difference

Equal variances assumed $\bar{x}_1 - \bar{x}_2 \pm t_{n_1+n_2-2; \alpha/2}\, S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$

$\sigma_1^2 = \sigma_2^2$ (άγνωστες αλλά ίσες)

Equal variances not assumed $\sigma_1^2 \neq \sigma_2^2$ και $n_1 = n_2$

δ.ε. $\left(\bar{x}_1 - \bar{x}_2 \pm t_{v;\alpha/2}\sqrt{\frac{s_1^2}{v} + \frac{s_2^2}{v}}\right)$, $v = 2(n-1)$

Σημειώστε ότι το $0 \in$ δ.ε. γεγονός που ενισχύει τις ενδείξεις ότι <u>δεν</u> υπάρχει σημαντική διαφορά στις μέσες τιμές στα δύο δείγματα.

<u>Εξαρτημένα δείγματα - Paired Samples</u> Πολλές φορές για την πραγματοποίηση μιας μελέτης είναι απαραίτητο να κάνουμε μετρήσεις πάνω στα <u>ίδια αντικείμενα</u> σε δύο διαφορετικές χρονικές στιγμές μετρώντας την ίδια παράμετρο (δηλαδή την ίδια μεταβλητή). Έστω ότι συγκρίνουμε δύο θεραπείες. Τα δείγματα στα οποία θα εφαρμοστούν οι δύο θεραπείες πρέπει να είναι όσον το δυνατόν πιο ομοιογενή. Αν επιτραπεί να μεταβληθούν κάποιες συνθήκες που επηρεάζουν τον τρόπο που αντιδρούν τα άτομα, μπορεί να μεταβληθούν κατά πολύ οι μετρήσεις και φυσικά και το αποτέλεσμα της μέτρησης. Η ανάγκη ομοιογένειας περιορίζει αρκετά τις δυνατότητες επιλογής των δειγμάτων. Για να συγκριθούν π.χ. δύο παυσίπονα θα πρέπει να επιλεχθούν δύο ομάδες ατόμων της ίδιας ηλικίας, η κατάσταση της υγείας τους να είναι η ίδια, να αντέχουν το ίδιο στον πόνο, πράγμα πολύ δύσκολο και ίσως πρακτικά ακατόρθωτο. Επιλέγεται μία ομάδα ατόμων και στα ίδια άτομα σε διαφορετικές χρονικές στιγμές δίνονται τα παυσίπονα και γίνονται οι μετρήσεις. Οι παρατηρήσεις που συγκεντρώνονται μ' αυτόν τον τρόπο καλούνται ζευγαρωτές παρατηρήσεις. Αν x_i, y_i είναι οι δύο μετρήσεις του i-οστού ατόμου, η x_i με την y_i δεν είναι ανεξάρτητες ενώ για διαφορετικά άτομα οι διαφορές $x_i - y_i$ για τα διάφορα i είναι ανεξάρτητες. Για κάθε i ορίζεται η διαφορά $d_i = x_i - y_i$, $i = 1, \ldots, n$ είναι ανεξάρτητες παρατηρήσεις, αποτελούν τ.δ. με μέση τιμή $\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$ και διασπορά

$$s_d^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} d_i^2 - n\bar{d}^2\right)$$

Ένα $100(1-\alpha)\%$ δ.ε. για τη διαφορά $\mu_1-\mu_2=\mu_d$ είναι το:

$$\left(\bar{d}-t_{n-1;\alpha/2}\frac{S_d}{\sqrt{n}},\ \bar{d}+t_{n-1;\alpha/2}\frac{S_d}{\sqrt{n}}\right)$$

για μικρά δείγματα, $n<30$ και με την προϋπόθεση ότι ο πληθυσμός από τον οποίον προέρχονται τα ζεύγη των παρατηρήσεων προέρχεται από την κανονική κατανομή.

Για $n\geq 30$ μπορούμε να γράψουμε: $t_{n-1;\alpha/2}\cong z_{\alpha/2}$.

<u>Παράδειγμα</u> Σε ένα πείραμα για τον εντοπισμό των πιθανών βλαβών στον ανθρώπινο εγκέφαλο από το αλκοόλ μία γνωστή τεχνική χρησιμοποιήθηκε για τη μέτρηση της πυκνότητας του εγκεφάλου σε 11 χρόνιους αλκοολικούς. Για κάθε αλκοολικό, ένας μη-αλκοολικός ίδιας ηλικίας, φύλου, εκπαίδευσης κι άλλων παραγόντων εντοπίστηκε και αναλύθηκε. Τα δεδομένα δίνονται στον πίνακα που ακολουθεί:

Αλκοολικός: 40.1 38.5 36.9 41.4 40.6 42.3 37.2 38.6 38.5 38.4 38.1

Μη-Αλκοολικός: 41.3 40.2 37.4 46.1 43.9 41.9 39.9 40.4 38.6 38.1 39.5

Το ενδιαφέρον μας αφορά την εύρεση διαφορών στη μέση πυκνότητα του εγκεφάλου μεταξύ αλκοολικών και μη.

Για να προχωρήσουμε στην εφαρμογή του t-test για ζευγαρωτές παρατηρήσεις θα πρέπει αρχικά να ελέγξουμε αν οι διαφορές των μετρήσεων $d_i=X_i-Y_i$ προέρχονται από πληθυσμό ο οποίος ακολουθεί την κανονική κατανομή. Κατί τέτοιο μπορεί να γίνει γραφικά με το Q-Q plot ή ακόμα ακριβέστερα με έλεγχο προσαρμογής της κανονικής κατανομής (π.χ. K-S test).

Δύο μεταβλητές → Alcool
Non Alcool

Transform → Compute Variable
Numeric Expression
Target Variable : D = ALCOOL − NONALCOOL

Για το Q-Q plot της μεταβλητής d έχουμε:

Analyze → Descriptive Statistics → Explore
Plots Normality plots with tests

Για το One-Sample Kolmogorov-Smirnov Test πάμε:

Analyze → Non-Parametric Tests → 1-Sample K-S

$\hat{\alpha}=0.519$ δεν απορρίπτεται η Ho της κανονικότητας.

Για το Paired-Samples T-test έχουμε

Analyze → Compare Means → Paired Samples T-test

Pair 1 → alcool - nonalcool

Αρχικά, λαμβάνουμε μία μικρή περιγραφική στατιστική των παρατηρήσεων (Paired Samples Statistics) και στη συνέχεια λαμβάνουμε τον συντελεστή συσχέτισης του Pearson για τη διαπίστωση της θεωρητικής ύπαρξης γραμμικής σχέσης μεταξύ των δύο μεταβλητών. Ο συντελεστής συσχέτισης του Pearson είναι υψηλός (0.797) υποδεικνύοντας την ύπαρξη μιας ισχυρής γραμμικής συσχέτισης μεταξύ των δύο μεταβλητών.
Ο συντελεστής συσχέτισης του Pearson είναι το πιο συχνά χρησιμοποιούμενο μέτρο συσχέτισης, συμβολίζεται με r και ορίζεται ως εξής: (βλέπε Ξεκαλάκη, σελ. 323-326).

$$r = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\left[\sum_{i=1}^{n}(X_i-\bar{X})^2\sum_{i=1}^{n}(Y_i-\bar{Y})^2\right]^{1/2}}$$

Η τιμή του μέτρου συσχέτισης r είναι μεταξύ -1 και $+1$, χαρακτηρίζοντας ανάλογα το είδος της συσχέτισης μεταξύ των μεταβλητών X και Y.

Η μέση τιμή των διαφορών στην πυκνότητα του εγκεφάλου μεταξύ αλκοολικών και μη ισούται με -1.5182 με αντίστοιχο 95% δ.ε. από -2.5795 έως -0.4569. Στο διάστημα αυτό δεν συμπεριλαμβάνεται το 0 και συνεπώς με πιθανότητα 5% τύπου I υπάρχει στατιστικά σημαντική διαφορά στη μέση πυκνότητα. Η τιμή του στατιστικού ελέγχου είναι $t=-3.187$ με 10 β.ε. και $\hat{\alpha}=0.01$. Το αρνητικό πρόσημο στην τιμή του στατιστικού ελέγχου μας οδηγεί να συμπεράνουμε επιπλέον ότι η μέση πυκνότητα των αλκοολικών είναι μικρότερη της αντίστοιχης των μη-αλκοολικών.

Παρατήρηση: Αν οι προϋποθέσεις για την εφαρμογή του paired t-test δεν ισχύουν (δηλ. οι διαφορές των μετρήσεων $d_i = X_i - Y_i$ **δεν** προέρχονται από πληθυσμό ο οποίος ακολουθεί την κανονική κατανομή) θα πρέπει να χρησιμοποιήσουμε κάποιο μη-παραμετρικό test. Συνηθέστερο είναι το Wilcoxon signed-rank test (έλεγχος των προσημασμένων τάξεων μεγέθους του Wilcoxon για δείγμα ζευγών παρατηρήσεων).

Θεωρούμε τις διαφορές $D_i = X_i - Y_i$, $i=1,\dots,n$ και στη συνέχεια λαμβάνουμε τις απόλυτες διαφορές $|D_i| = |X_i - Y_i|$, $i=1,\dots,n$ αγνοώντας τα ζεύγη (X_i, Y_i) για τα οποία $X_i = Y_i$ δηλαδή $D_i = 0$. Μετά διατάσσουμε κατά αύξουσα τάξη μεγέθους τις n' ($n' \leq n$) διαφορές $|D_i|$ που απομένουν.

Βαθμός 1: Μικρότερη απόλυτη διαφορά $|D_i|$

⋮

Βαθμός n': Μεγαλύτερη απόλυτη διαφορά $|D_i|$

Αν αρκετά ζεύγη έχουν απόλυτες διαφορές που ταυτίζονται τότε σε καθένα από αυτά τα ζεύγη αντιστοιχίζεται ως βαθμός ο μέσος των βαθμών που τα ζεύγη θα είχαν αν οι διαφορές τους δεν ήταν ίσες. Έστω $d_{0.5}$ η διάμεσος του πληθυσμού των διαφορών. Ελέγχουμε: A. $H_0: d_{0.5} \le 0$, $H_1: d_{0.5} > 0$ B. $H_0: d_{0.5} \ge 0$, $H_1: d_{0.5} < 0$

Γ. $H_0: d_{0.5} = 0$
$H_1: d_{0.5} \neq 0$

Ορίζουμε $R_i = \begin{cases} +R(|D_i|), & \text{αν } D_i = X_i - Y_i > 0 \\ -R(|D_i|), & \text{αν } D_i = X_i - Y_i < 0 \end{cases}$ $i = 1, \dots, n'$

Στατ. συνάρτηση ελέγχου: $T = \dfrac{\sum_{i=1}^{n'} R_i}{\sqrt{\sum_{i=1}^{n'} R_i^2}}$

Αν όλες οι διαφορές είναι διακεκριμμένες, εναλλακτικά

$$T = \frac{\sum_{i=1}^{n'} R_i}{\sqrt{n(n+1)(2n+1)/6}}$$

ή T^+:= άθροισμα των θετικών R_i

όπου R_i ο βαθμός των διαφορών $|D_i|$, $i = 1, 2, \dots, n'$

Κρίσιμες περιοχές A. $T > z_{1-\alpha}$ ή $T^+ > w_{1-\alpha}$ (W: ΠΙΝΑΚΑΣ 8 ΞΕΚΑΛΑΚΗ)

B. $T < z_{\alpha} = -z_{1-\alpha}$
ή $T^+ < w_{\alpha}$

Γ. $T < z_{\alpha/2}$ και $T > z_{1-\alpha/2}$ ή $T^+ < w_{\alpha/2}$ και $T^+ > w_{1-\alpha/2}$

Στο παράδειγμά μας πάμε: Analyze → Non-Parametric Tests → Two Related-Samples-Test

$\hat{\alpha} = 0.013 \qquad \alpha = 0.05 > \hat{\alpha}$

Υπάρχει στατιστικά σημαντική διαφορά στη μέση πυκνότητα του εγκεφάλου μεταξύ αλκοολικών και μη.

Τα άλλα τρία test (Sign Test, McNemar και Marginal Homogeneity Test) είναι εναλλακτικά test, του Wilcoxon. Το McNemar Test χρησιμοποιείται όταν θέλουμε να ελέγξουμε τη συμπεριφορά των ατόμων ενός πληθυσμού και τη μεταβολή της, πριν και μετά τη μεσολάβηση κάποιου γεγονότος. Μία επέκταση του McNemar Test που επιτρέπει περισσότερες από δύο επιλογές (πριν και μετά) είναι το Marginal Homogeneity Test. Το Sign Test (έλεγχος προσήμων) χρησιμοποιείται κυρίως για να ελέγξει αν οι τιμές μίας από τις τ.μ. του ζεύγους (X,Y) τείνουν να είναι μεγαλύτερες ή μικρότερες από τις τιμές της άλλης (βλέπε ΞΕΚΑΛΑΚΗ, σελ. 20).

Έλεγχος McNemar: Το έχουμε αναφέρει στο ΚΕΦ. 2 ΣΤΑΤΙΣΤΙΚΗ ΣΥΜΠΕΡΑΣΜΑΤΟΛΟΓΙΑ ΓΙΑ ΤΗΝ ΑΝΑΛΥΣΗ ΜΙΑΣ ΜΕΤΑΒΛΗΤΗΣ.