$$\hat{p}_1 - \hat{p}_2 \sim N\left(0,\ p(1-p)\left(\frac{1}{n} + \frac{1}{m}\right)\right)$$

Η παράμετρος όμως p είναι άγνωστη και πρέπει να εκτιμηθεί από τα στοιχεία των δειγμάτων που έχουμε. Η κοινή αναλογία των επιτυχιών σ' ένα δείγμα μεγέθους n + m το $\hat{p} = \dfrac{x+y}{n+m}$, είναι μία εκτίμηση του p. Αν τώρα τα n και m είναι μεγάλα ισχύει:

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{p(1-p)\left(\frac{1}{n} + \frac{1}{m}\right)}} \sim N(0,1)$$

και το test για τη σύγκριση δύο αναλογιών p_1 και p_2 είναι αυτό που περιγράψαμε στην παράγραφο 6.7 δηλαδή

H_0: $\quad p_1 = p_2$
H_1:
1. $p_1 > p_2$
2. $p_1 < p_2$
3. $p_1 \neq p_2$

Απορρίπτουμε την H_0 αν

1. $z > Z_a$
2. $z < -Z_a$ όπου $z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n} + \frac{1}{m}\right)}}$
3. $z > Z_{a/2}$

$$\hat{p}_1 = \frac{x}{n}, \quad \hat{p}_2 = \frac{y}{m}, \quad p = \frac{x+y}{n+m} \quad \text{και} \quad n, m \geqslant 30.$$

Τέλος Z_a και $Z_{a/2}$ οι γνωστές τιμές της τυπικής κατανομής.

6.11. Σχέση ανάμεσα στα test σημαντικότητας και στα διαστήματα εμπιστοσύνης.

Σ' αυτό το σημείο είναι χρήσιμο να παρατηρήσουμε ότι ανάμεσα στους τύπους που αναφέραμε στο κεφάλαιο 5 και στους τύπους σ' αυτό το κεφάλαιο, υπάρχει μια σχετική ομοιότητα. Π.χ. ένα $100(1-a)\%$ δ.ε. για τη μέση τιμή μ του πληθυσμού δίνεται από τη σχέση

$$\bar{x} - t_{n-1, a/2} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{n-1; a/2} \frac{s}{\sqrt{n}} \tag{6.3}$$

Από την άλλη πλευρά η απορριπτική περιοχή για το δίπλευρο test

H_0: $\mu = \mu_0$

H_1: $\mu \neq \mu_0$ είναι $R: \left| \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} \right| \geqslant t_{n-1; a/2}$.

αυτό σημαίνει ότι για τις τιμές $\left| \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} \right| < t_{n-1; a/2}$ δεχόμαστε

την μηδενική υπόθεση. Η παραπάνω ανισότητα γράφεται:

$$\bar{x} - t_{n-1; a/2} \frac{s}{\sqrt{n}} < \mu_0 < \bar{x} + t_{n-1; a/2} \frac{s}{\sqrt{n}} \tag{6.4}$$

Οι τύποι 6.3 και 6.4 συνδυαζόμενοι λένε το εξής: δεχόμαστε την H_0 σε σ.σ. a όταν η τιμή μ_0 ανήκει στο $100\,(1-a)\%$ δ.ε. της μέσης τιμής του πληθυσμού.

Με απλά λόγια, πολλές φορές για να ελέγξουμε την H_0: $\theta = \theta_0$ (όπου θ μπορεί να είναι μία από τις παραμέτρους μ, σ, p), βρίσκουμε το δ.ε. για την ίδια σ.σ. και για την ίδια παράμετρο που ελέγχουμε. Αν τώρα η τιμή θ_0, βρίσκεται μέσα στο δ.ε., τότε δεχόμαστε την H_0, αλλοιώς την απορρίπτουμε. Ο τρόπος αυτός, επειδή μπορεί να ελέγξει συγχρόνως πολλές μηδενικές υποθέσεις, θεωρείται αρκετά χρήσιμος στη στατιστική.

6.12. Μέγεθος δείγματος.

Στους τύπους που αναφέραμε στις προηγούμενες παραγράφους, όπως στον υπολογισμό του ανώτατου ορίου του λάθους στην εκτίμηση της μέσης τιμής ή στον υπολογισμό του δ.ε. πολλές φορές εμφανίσθηκε η ποσότητα $\sqrt{n}$ στον παρονομαστή.