1. X – просечна оцена атлетата који су похађали спец. тренинг

Y – –||– који нису похађали специјални тренинг

– потребно је применити Вилкоксонов тест за 2 независна узорка $(X_1,\dots,X_n)$ и $(Y_1,\dots,Y_m)$, $n=11$, $m=12$

– претпоставка овог теста: $X \stackrel{d}{=} Y+c$

– ако атлетичари који су похађали спец. тренинг надмашују у резултатима остале атлетичаре, важи $c>0$ (или еквивалентно томе: $EX > EY$, јер $X \stackrel{d}{=} Y+c \Rightarrow EX=EY+c$)

Дакле, тестираћемо следеће хипотезе:

$$H_0: c=0 \quad \text{vs} \quad H_1: c>0$$

– тест стат: $T=\sum_{i=1}^{n} R_i$, где је R_i ранг од X_i у обједињеном узорку $X_1,\dots,X_n,Y_1,\dots,Y_m$

– израчунајмо реализовану вредност тест стат.:

1. сортирамо обједињени узорак
2. рангирамо га
3. саберемо рангове X-ева

0	1	2	2.5	3.1	4.3	5.1	5.5	6	6.3	7	7	7.2	7.9	8	8.1	8.4	9	9	9	9.1	9.1	9.7
1	2	3	4	5	6	7	8	9	10	11.5	11.5	13	14	15	16	17	19	19	19	21.5	21.5	23
				X	X					X		X	X	X	X	X	X	X		X		

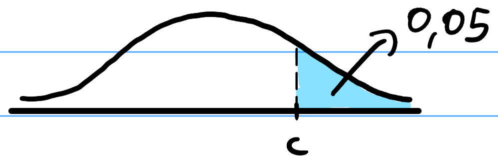
$$\Rightarrow t = 5+6+11.5+13+14+15+16+17+19+19+21.5 = 175$$

- $n, m > 10$, можемо користити стандардизовану тест статистику:

$$T^* = \frac{T - \frac{n(n+m+1)}{2}}{\sqrt{\frac{nm(n+m+1)}{12}}} \underset{\text{при } H_0}{\sim} \mathcal{N}(0,1)$$

- критична област $W = \{T^* \geq c\}$,

c је т.д. $P\{T^* \geq c \mid H_0 \text{ тачна}\} = 0.05$



$$\Rightarrow \Phi(c) = 0{,}95 \Rightarrow c = \Phi^{-1}(0{,}95) = 1.64$$

- реализована вредност T^*:

$$t^* = \frac{175 - \frac{11(11+12+1)}{2}}{\sqrt{\frac{11 \cdot 12(11+12+1)}{12}}} = 2.64$$

- $t^* \geq c$, па одбацујемо H_0 и прихватамо тврдњу компаније

(2.) Играч жели да тестира следеће хипотезе

$H_0: P\{\text{победа}\} = P\{\text{пораз}\} = P\{\text{нерешено}\} = \frac{1}{3}$

vs H_1: не важи H_0

- Нека исходи партија буду наше категорије:

	победа	пораз	нерешено
m_k	4	13	7

$k = 3$, $n = 24$

- тест стат: $T = \sum_{j=1}^{k} \frac{(M_j - np_j)^2}{np_j} \underset{\text{при } H_0}{\sim} \chi^2_{k-1}$,

где је $p_j = P\{j\text{-ти исход партије} \mid H_0 \text{ тачна}\} = \frac{1}{3}$, $j \in \{1,2,3\}$

- провера да су категорије довољно велике:
$np_j = 8 > 5$, $j \in \{1,2,3\}$

а) Како је критична област овог теста облика:
$W = \{T \geq c\}$,

p - вредност теста је:

$p = P\{T \geq t \mid H_0 \text{ тачна}\}$

$$p = 1 - F_{\chi^2_2}(t),$$ где је t реализована вредност тест стат. T

$$t = \frac{(4-8)^2}{8} + \frac{(13-8)^2}{8} + \frac{(7-8)^2}{8} = 5.25$$

$\Rightarrow p \approx 0.07$ (у R-у: $F_{\chi^2_2}(5.25) =$ pchisq(5.25, 2))

б) како је p-вредност најмањи ниво значајности за који ћемо одбацити H_0 на основу датог узорка, следи да због $0.05 < p$ не бисмо одбацили H_0 за ниво значајности 0.05

3. - тражи нам се p-вредност χ^2 теста независности пола и омиљене боје ученика

	плава	зелена	розе
дечаци	100	150	20
девојчице	20	30	180

H_0: пол и боја независни
vs
H_1: пол и боја зависни

$\Rightarrow$ укупан бр. дечака $= 100 + 150 + 20 = 270$
укупан бр. девојчица $= 20 + 30 + 180 = 230$
$n = 270 + 230$ (укупан бр. ученика) $= 500$

- тест стат:

$$T = \frac{(M_{mp} - n\hat{p}_{m\cdot}\hat{p}_{\cdot p})^2}{n\hat{p}_{m\cdot}\hat{p}_{\cdot p}} + \frac{(M_{mz} - n\hat{p}_{m\cdot}\cdot\hat{p}_{\cdot z})^2}{n\hat{p}_{m\cdot}\cdot\hat{p}_{\cdot z}} + \dots + \frac{(M_{žr} - n\hat{p}_{ž\cdot}\hat{p}_{\cdot r})^2}{n\cdot\hat{p}_{ž\cdot}\hat{p}_{\cdot r}},$$

Где је: M_{mp} - број дечака чија је омиљена боја плава,
M_{mz} - број дечака чија је омиљена боја зелена,
⋮
$M_{žr}$ - број девојчица чија је омиљена боја розе.

$\hat{p}_{m\cdot}$ - оцена вероватноће да је изабрани ученик дечак
$\hat{p}_{ž\cdot}$ - -||- девојчица
$\hat{p}_{\cdot p}$ - оцена вероватноће да је омиљена боја ученика плава
$\hat{p}_{\cdot z}$ - -||- зелена
$\hat{p}_{\cdot r}$ - -||- розе

При H_0 $\quad T \sim \chi^2_{(3-1)(2-1)}$

- да бисмо израчунали реализовану вредност од T, треба да израчунамо све оцене вероватноћа које се помињу:

нпр. $\hat{P}_{m\cdot} = \dfrac{\text{број дечака}}{\text{број ученика}} = \dfrac{270}{500}$

слично: $\hat{P}_{ž\cdot} = \dfrac{230}{500}$

$$\hat{P}_{\cdot p} = \frac{\text{бр. ученика чија је омиљена боја плава}}{\text{бр. ученика}} = \frac{120}{500}$$

$$\hat{P}_{\cdot z} = \frac{180}{500}$$

$$\hat{P}_{\cdot r} = \frac{200}{500}$$

- реал. вр. од T:

$$t = \frac{\left(100 - 500\cdot\frac{270}{500}\cdot\frac{230}{500}\right)^2}{500\cdot\frac{270}{500}\cdot\frac{230}{500}} + \ldots + \frac{\left(180 - 500\cdot\frac{230}{500}\cdot\frac{200}{500}\right)^2}{500\cdot\frac{230}{500}\cdot\frac{200}{500}} = 259.76$$

– критична област: $W = \{T \geq c\}$, па ћемо p-вредност наћи као:

$$p = P\{T \geq t \mid H_0 \text{ тачна}\}$$

$$= 1 - F_{\chi^2_2}(t) \approx 0$$

– мала p-вредност сугерише на то да би требало одбацити H_0, тј. прихватити хипотезу о постојању зависности између пола ученика и њихове омиљене боје.