پویا شیخ الاسلامی - تکلیف کد سوم
۸۱۰۱۰۰۳۹۴ - استنباط آماری

a)
① I
بازی را به گونه‌ای طراحی می‌کنم که میمون‌ها (players / Agents) در طی یک فاصله زمانی بتوانند انتخاب‌های خود و نتایج آن را ببینند.
همچنین از طریق هر وسیله‌ای بتوانند (مانند Keyboard و ...) استراتژی خود را انتخاب کنند (▲ و ●) و سپس در با این براساس انتخاب‌ها به آنها پاداش داده شود و این روند برای هر جفت Agent (میمون) n بار تکرار می‌شود.

		$\triangle$ P_2	$\bigcirc$
P_1	$\triangle$	a,a	c,d
	$\bigcirc$	d,c	b,b

(*)

I
۱) برای حالت نخست (فرض H_0) میمون‌ها در صورت کمبود به صورت صحیح می‌توانند در نظر گرفته شود یعنی بازی مثلاً از یک گروه میمون گرفته شود و نتایج آن از نظر پاداش با هم مقایسه شود (میزان پاداش گرفته شده / پاسخ گویی صحیح) اگر تفاوت بین دو گروه نباشد اگر تفاوت معنادار نباشد H_0 رد نمی‌شود و اگر تفاوت کنند H_1 (فرض اولیه) رد می‌شود.
۲) برای حالت دوم (H_0) میمون‌ها علائم تطابقی مثل زمان پاسخ‌گویی و یا علائم ناخواسته و علائم فیزیکی مثل تپش قلب و ... را در نظر گرفت که در n تکرار این بین دو گروه این علائم مشابه شوند اگر تفاوت کنند H_0 رد شود.....
II از تست non-parametric باید استفاده شود. چون نمی‌توانیم از فرضیاتی در مورد داده‌ها استفاده کنیم مثلاً تست ANOVA می‌توانند مناسب‌تر که از طرح‌های بررسی جلوگیری می‌کنند و هر دو فرض را به طور هم زمان مقایسه می‌کنند.

III توزیع پاداش برای هر میمون باید به دلیل اثر یادگیری میمون و همچنین تأثیر خستگی بر روند ذهنی یک توزیع نرمال با چولگی به چپ باشد.

دلیل چولگی - راست همان فرایند یادگیری است که باید مقداری طول بکشد و دلیل چولگی چپ هم تأثیر خستگی ذهن در روند انجام آزمایش و دریافت پاداش است. همچنین برای تعیین پاداش (توزیع) باید نتایج n بار تکرار را برای هر زوج میمون بد (n تعداد تکرار) نرمالیزه کنیم.
IV برای اختلاف معناداری بین میمون‌ها اگر پاداش را براساس پاسخ‌های درست داده شده بازی بسیار رقابتی می‌شود پس باید پاداش - میمون داده شود که به طور عادلانه پاسخ‌ها را با دلیل میمون‌ها دارد (و نه لزوماً پاسخ درست) طبق شکل بازی (*) مقادیر روی قطر اصلی را زیاد و قطر فرعی را کم

b
I می‌توان برای مشاهده تفاوت تحمل و همان تفسیر استراتژی توابع سود (utility) را در نظر گرفت و مقایسه کرد که این تفسیر استراتژی چطور و در چه جهتی تغییر کرده‌اند. Wilcoxon signed rank
می‌توان چون جمعیت - اندازه‌ی کافی بزرگ داریم (فرض ۱) از z stat استفاده کرد تا ببینیم که تفسیر استراتژی در چه جهتی صورت گرفته است. (بین male / female)
II می‌توانیم فرضیات اولیه انجام دهیم پس با parametric test می‌توان استفاده کرد مثلاً بین تعداد کل میمون‌ها (70) تست Mann-Whitney که طرح فرضیه کرد و مشاهده کرد که آیا تفاوت معناداری بین پاداش‌ها وجود دارد یا خیر... (بین دو گروه male / female)
III می‌توان از تست chi-square استفاده کرد و اختلاف بین میمون‌ها را با سال سن را در دو گروه مقایسه کرد و معلوم کرد که تفاوت معنادار دارد یا خیر
در حقیقت سن بلوغ ملاک میمون‌ها را در اینجا به در نظر گرفته شده است که حال باید ببینیم که این دو گروه با هم در پاداش تفاوتی دارند یا خیر (male / female)

(c) در این نوع تست friedman test می‌کوشند که با استفاده از یک ابزار (n مرتبه) آزمون برای همه‌ی Agentهای تولید بین عملکرد آنها تفاوت گذاشت و آن‌ها را مرتب و رتبه‌بندی کرد (بر اساس شاخص‌های گوناگون)

(2)

a) $E = n \cdot p = 100(0.5) = 50$ $\qquad df = 2-1 = \boxed{1}$

$$\chi^2 = \sum_{i=1}^{2} \frac{(x_i - \mu_i)^2}{\mu_i} = \frac{(38-50)^2}{50} + \frac{(62-50)^2}{50} = \boxed{5.76}$$

$H_0: p_1 = p_2 = 0.5$

$$p(x \leq 38) \simeq p\left(z < \frac{38-50}{5}\right) \simeq \boxed{0.0081} \xrightarrow{\times 2} \text{two-tailed} = 0.0163$$

(2)

b) طبق توضیحات چون از chi square با درجه آزادی χ^2 استفاده می‌کنیم $df = 1$ پیروی کند
$2-1=1$

c) با توجه به آنچه در قسمت قبل بدست آوردیم $df = 1$ را از جدول مربوطه جایگذاری می‌کنیم داریم:

df	0.90	0.95	0.975	0.99	0.995
1	2.71	3.84	5.02	6.63	9.14
2	4.61	5.99	7.38	9.21	11.28
3	6.25	7.81	9.35	11.34	14.32
4	7.78	9.44	11.14	13.28	16.42

$1 - V_{al}(9.99) < Pvalue < 1 - V_{al}(9.975)$

$0.01 < Pvalue < 0.025$

d) $\dfrac{(x-\frac{n}{2})^2}{\frac{n}{2}} + \dfrac{(n-x-\frac{n}{2})^2}{\frac{n}{2}}$ تابع یک تابع درجه دو است و کمترین مقدار آن برابر صفر است $(x=\frac{n}{2})$ تابع یک تابع صعودی است

$\frac{2}{n}\left(\left[(x-\frac{n}{2})^2\right] + (x-\frac{n}{2})^2\right) = \frac{4}{n}(x-\frac{n}{2})^2$

$|x_1-\frac{n}{2}| > |x_2-\frac{n}{2}| \Rightarrow$ و با افزایش n نیز مقدارش افزایش می‌یابد دلیل هم
$(x_1-\frac{n}{2})^2 > (x_2-\frac{n}{2})^2$

احتمال
$x = x \longrightarrow \binom{n}{x}(05)^n$

e) $n \gg 0$ (تعداد بسیار بزرگ) شبیه به فرایند پواسون
چون مقایسه بزرگ است (به بی‌نهایت محدود) می‌توان فرض کرد با مدل توزیع پواسون طرف هستیم

$$\begin{cases} p(X=x) = \dfrac{\lambda^x}{x!} \cdot e^{-\lambda} \\ \lambda = \frac{n}{2} \end{cases}$$

$\longrightarrow |x-\frac{n}{2}| > k \longrightarrow x < \frac{n}{2} - k$
$|x-\frac{n}{2}| > 0$, $k > 0$

$p(x) < \dfrac{(\frac{n}{2})^{\frac{n}{2}-k}}{(\frac{n}{2}-k)!} e^{-\frac{n}{2}} \Longrightarrow$

$S.L = \dfrac{(\frac{n}{2})^{[\frac{n}{2}-k]}}{[\frac{n}{2}-k]!} e^{-\frac{n}{2}}$

4

a) فرض صفر H_0: توزیع تولد بازیکنان لیگ برتر $=$ توزیع تولد مردم ایران
فرض Alternative H_1: $\neq$

b) (I) تمام نمونه‌ها باید مستقل باشند (باید بر اساس (simple random sampling) نمونه برداری شوند)
(II) متغیر باید از نوع دسته‌بندی (catagorial) باشد.
(III) Expected value برای هر شاخه (تعداد هر متغیر) در هر طبقه باید دست کم ≥ 5 باشد.

c) $\chi_c^2 = \sum \dfrac{(x_i - np_i(\hat{\theta}))^2}{np_i(\hat{\theta})}$

$n = 147 + 110 + 52 + 50 = \boxed{359}$

$\nu = k - m - 1 = 4 - 1 - 0 = \boxed{3}$ درجه آزادی

$\chi_3^2 = \dfrac{(147-85)^2}{85} + \dots \Rightarrow$

$\chi_3^2 = 12.48 + 0.85 + 5.10 + 4.85 = \boxed{82.6}$

$\boxed{Pvalue \simeq 0}$ مقدار Pvalue بسیار ناچیز خواهد شد.

R (d

c) بر اساس Pvalue به دست آمده می توان نتیجه گرفت که فرض صفر (H_0) رد می شود از برنامه R داریم $\text{Pvalue} = 0.009032$

چون مقدار اطمینان است نفی فرض اولیه با شانس کمی رخ خواهد داد رد خواهد شد.

تست دوطرفه

CE_IQ	125	127	126	126	105	128	127	126	79	[illegible]
rank	9	13.5	11	11	7	15	13.5	11	5	8
EE_IQ	131	129	77	52	134	133	91	68	67	132
rank	17	16	4	1	20	19	6	3	2	18

a)

H_0: IQ score دانشجویان برق و کامپیوتر معادل یکدیگرند.

H_1: IQ score دانشجویان برق و کامپیوتر معادل نیستند.

b) ابتدا باید داده ها را مرتب کنیم و با توجه به ترتیب دو گروه (مشترک) آن ها را رنک کنیم.

$$R_1 = 4 + 13.5 + 11 + \dots = \boxed{104} \quad (EE)$$

$$R_2 = 17 + 16 + \dots = \boxed{106} \quad (CE)$$

c) $$U_i = n_1 n_2 + \frac{n_i(n_i+1)}{2} - R_i$$

$$\begin{cases} U_1 = 100 + 55 - 104 = 51 \ (EE) \\ U_2 = 100 + 55 - 106 = 49 \ (CE) \end{cases} \longrightarrow U = \min(U_1, U_2) = \boxed{49}$$

با توجه به مقایسه در تست دوطرفه مقدار Pvalue 2٪ استراتژی تمام Pvalue ها که کوچکتر از 49 که کوچک است.) پس فرض H_0 رد نمی شود و با شانس خوبی درست است.

d) $P \simeq 0.97$ (از R)

e) به طبق آن چه به دست آمد با احتمال 96.8٪ فرض H_0 برقرار است و قابل رد کردن نیست. نفی در گروه معادل یکدیگرند

(b

(a به دلیل iid بودن داریم

$$p = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}} = \frac{1}{(\sigma\sqrt{2\pi})^n} \prod_{1}^{n} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}} \xrightarrow{\ln}$$

برای max

$$\ln p = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{1}^{n}(x_i-\mu)^2 \xRightarrow[\frac{\partial}{\partial \sigma^2}\ln p = 0]{\frac{\partial}{\partial \mu}\ln p = 0}$$

$$\begin{cases} \frac{2}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = \frac{1}{\sigma^2}\sum(x_i-\mu) = 0 \\ \frac{-n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 = 0 \end{cases} \Rightarrow \begin{cases} \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i \\ \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2 \end{cases}$$

b) ویژگی استاندارد

$$S = \sqrt{n-1}\,\frac{\hat{\mu}-\mu}{\sqrt{\hat{\sigma}^2}} = \frac{\hat{\mu}-\mu}{\frac{\hat{\sigma}}{\sqrt{n-1}}} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i - \mu}{\frac{1}{\sqrt{n(n-1)}}\sqrt{\sum_{i=1}^{n}(x_i-\mu)^2}} = \frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{S/\sigma} = \frac{\sqrt{n}(\bar{X}-\mu)}{S} \sim t_{n-1}$$

$Z = dN(0,1)$ $Z = dN(0,1)$

c) $N(0,1)$ به $\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$ برای nهای بزرگ $n \gg$.

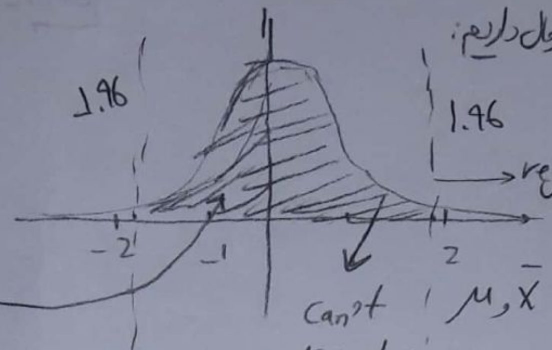
برای یک متغیر تصادفی با توزیع نرمال داریم:

$$P_r\{-1.96 < Z < 1.96\} = \boxed{0.95}$$

$$Z \downarrow \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$$

$$\downarrow \div \frac{\sigma}{\sqrt{n}}$$

$$P_r\{-1.96\frac{\sigma}{\sqrt{n}} < \bar{x}-\mu < 1.96\frac{\sigma}{\sqrt{n}}\} \simeq \boxed{0.95}$$



$\boxed{\alpha = 5\%}$

یعنی با احتمال ۹۵٪ اختلاف μ, $\bar{x}$ کمتر از $1.96\frac{\sigma}{\sqrt{n}}$ است.

(نمی‌توانیم فرض صفر را رد کنیم.)

---

7)

a) چون wilcoxon signed rank برای دو مجموعه داده جفت (paired) است.

از Mean-Whitney یا rank sum استفاده می‌کنیم:

rank	13	7	11	9	8	2	6	3.5
group1	19.0	14.4	18.2	15.6	14.5	11.2	13.9	11.6
group2	12.1	19.1	11.6	21.0	16.7	10.1	18.3	20.5
rank	5	14	3.5	16	10	1	12	15

دوطرفه استفاده می‌کنیم:

$$\begin{cases} S_1 = \sum^{group1} rank = \boxed{59.5} \\ S_2 = \sum^{g_2} rank = \boxed{76.5} \end{cases}$$

H_0: دو گروه تفاوتی از نظر تست واکنش ندارند.

H_1: دو گروه از نظر تست واکنش متفاوت‌اند.

$$\Rightarrow \begin{cases} U_1 = 64 + 36 - 59.5 = \boxed{40.5} \\ U_2 = 64 + 36 - 76.5 = \boxed{23.5} \end{cases} \longrightarrow U = \min(U_1, U_2) = 23.5 \longrightarrow \begin{cases} \alpha = 5\% \to U = 13 \\ \alpha = 1\% \to U = 30 \end{cases}$$

با دقت ۵٪ فرض H_0 غیرقابل رد کردن و با فرض ۱٪ = α H_0 می‌تواند رد شود.

b) مطابق روش تست رنک‌ها را انجام می‌دهیم

rank	7	13	11	15	16	4	12	6.5
g_1	19	14.4	18.2	15.6	14.5	11.2	13.9	11.6
g_2	12.1	19.1	11.6	21.0	16.7	10.1	18.3	20.5
rank	9	6	6.5	2	14	1	10	3

داریم

$$\begin{cases} S_1 = 84.5,\ S_2 = 51.5 \\ U_1 = 15.5 \\ U_2 = 48.5 \end{cases} \longrightarrow U = \boxed{16.5}$$

از روی جدول می‌دانیم که از مقدار بحرانی (13) با دقت ۵٪ بالاتر است

پس با این دقت (۵٪) فرض H_0 غیرقابل رد کردن است و با دقت ۱٪ = α

H_0 قابل رد کردن است.

---

c) در تست smooth به جای توزیع از تابع احتمال تجمعی (CDF) استفاده می‌شود. مهم‌ترین مزیت‌ها این است عبارتند از:

① در صورت رد کردن H_0 می‌تواند فرض جایگزین قابل قبولی را برخلاف تست‌های دیگر ارائه دهد.

۲) از توابع متعامد برای توصیف استفاده می‌کنند که هر توزیعی را می‌توان بر اساس این پایه‌ها نشان داد.

$$h(y) = C(\theta)\exp\left[\sum_{i=1}^{k}\theta_i \pi_i(y)\right] \longrightarrow \begin{cases} H_0: \theta_1 = \dots = \theta_n = \theta \\ H_1: \forall i,j \rightarrow \theta_i \neq \theta_j \end{cases}$$

ضریب ثابت

normalizer

توابع متعامد

دانش اینکه بتوان توابع متعامد را به صورت مثال به شکل زیر به دست آورد. ...

g_1	19	14.4	18.2	15.6	14.5	11.2	13.9	11.6
g_2	12.1	19.1	11.6	21	16.7	10.1	18.3	20.5
$\frac{g_1+g_2}{2} = \varphi_1$	15.55	16.75	14.9	18.3	15.6	10.65	16.1	16.05
$\frac{\lvert g_1-g_2\rvert}{2} = \varphi_2$	3.45	2.35	3.3	2.7	1.1	0.55	2.2	4.45

داریم

(الحاقی جاییست)

۸)

a) فرض:
H_0: طرح‌های بازاریابی تفاوتی با هم ندارند
H_1: با هم فرق می‌کنند ...
(اثر دارند یا نه)
(با استفاده از تست wilcoxon signed rank)

before	60	56	80	73	~~23~~ 14	32
after	58	58	83	67	~~67~~ 17	36
diff	-2	+2	+3	-6	+3	+4
rank	1.5	1.5	3.5	6	3.5	5

$T^+ = 1.5 + \dots + 5 = 13.5$

$T^- = 1.5 + 6 = 7.5$

$\longrightarrow T = \min(T^+, T^-) = \boxed{7.5}$

چون فرض ما از نوع افزایش است باید از تست همبسته (یک طرفه) استفاده کنیم به مقدار α را دو برابر کنیم و تست دو طرفه را اجرا کنیم.

دقت

برای $\alpha = 5\%$ (دقت تک طرفه) ← $\alpha = 10\%$ (دو طرفه)

b) کد ← برای این مقدار فرض H_0 قابل رد کردن نیست و تعداد طرح‌ها هر برنامه افزایش معنادار نداشته‌اند.

c)

$\mu_{dif} = \frac{102}{9} = \boxed{11.33}$

$S_D = \sqrt{\frac{\sum(x_i - \bar{x}_i)^2}{n-1}} = \boxed{13.02}$

%	A	B	C	D	E	F	G	H	I
before	78	50	40	40	20	50	50	50	50
after	65	20	50	8	16	44	38	40	45
diff	13	30	-10	32	4	6	12	10	5

$$\begin{cases} H_0: \mu_x \geq \mu_y \ (\mu_{dif} \geq 0) \\ H_1: \mu_x < \mu_y \ (\mu_{dif} < 0) \end{cases}$$

$t_0 = \frac{11.33}{\frac{13.02}{3}} = \boxed{2.61}$

$t_{1-\alpha}(n-1) = t_{1-0.05}(9-1) = \boxed{1.86}$

$t_0 = 2.61 > 1.86 \rightarrow$ پس نمی‌توانیم H_0 را رد کنیم.

۹) اما با توجه به آنچه که گفته شد چون روش صحیح نیست (چون نمونه برداری به صورت کاملاً تصادفی انجام نمی شود) بنابراین مسئله تست x^2 که اختلاف نسبت ها را مقایسه می کند نتیجه‌گیری غلط است. مغالطه ای که در اینجا موجود است این است که بدون توجه به ترکیب جمعی و فقط با در نظر گرفتن جمعیت بستری می بینیم که نسبت افراد مبتلا به هر دو بیماری، نسبت به افراد دارای یک بیماری، بیشتر از نسبت افراد مبتلا به بیماری اول به افراد سالم است. اما اگر همه حالات [هر ۴ حالت ممکن] را در نظر بگیریم این مشکل وجود ندارد.

۱۰) در این مغالطه، پارادوکس brekson می گویند که عدم وجود همبستگی را به possitive / negetive بودن نشان میدهند در دلایلشان مشاهدات نادرست است.
مثلی
نمی انتخاب جامعه آماری به طور random صورت نگرفته است. در اینجا بیماری وقتی نسبت ها می شود که مثلاً از نظر بدنی شان احتیاج به بیمارستان داشته باشند که همچنین موضوعی در محاسبات و مشاهدات وارد شده است که همین امر موجب بروز مشکل و ایجاد آن شده است.