امیرحسین محمدی 810197289

9 | 1

(أ) ابتدا باید آزمایشی برای مشارکت دو میمون طراحی کرد. آزمون به شکل زیر است به میمون ها ابزاری برای انتخاب بین دو شکل (در مثال ما دایره و مثلث) می‌دهیم. حال براساس انتخاب دو میمون شروع به جایزه دادن به میمون ها می‌کنیم و بعد از هر مرحله انتخاب میمون دیگر را به میمون مورد نظر نمایش می‌دهیم تا بتواند رفتار میمون دیگر را در دور بعدی پیش بینی کند. جایزه میمون ها بسیار شبیه معمای زندانی است به این نحو اگر میمون یک همکاری کند اما میمون دو نکند میمون دو 4 واحد آب میوه جایزه و میمون یک 0 واحد و برعکس. اگر هر دو همکاری کنند هرکدام 3 واحد و اگر هیچ کدام همکاری نکنند هر دو 1 واحد جایزه می‌گیرند. جدول آن به شرح زیر می‌باشد که همکاری متناظر با انتخاب دایره و عدم همکاری متناظر با مثلث است.

میمون 1 \ میمون 2	C	D
C	(3 , 3)	(0 , 4)
D	(4 , 0)	(1 , 1)

برای تست نیاز به دنبال جواب دادن به این سوال هستیم که آیا درک و واکنش احساس گونه های مختلف به پارامتر های مشخص متفاوت یکسان است یا خیر و همچنین آنکه رفتار و ذهن (همچنین تعداد نورون ها) پاسخ مشابهی جلوی پارامتر ها دارند یا خیر.

برای پاسخ به سوالات بالا دو تست آماری در نظر می‌گیریم.

تست اول:

H_0 : درک و پاسخ احساسی گونه ها به پارامتر های یادشده یکسان است.

H_1 : درک و پاسخ احساسی گونه ها به پارامتر های یادشده متفاوت است.

تست دوم:

H_0 : پاسخ های رفتاری و روان شناختی از جمله هورمون ها و نورون ها به پارامتر ها یکسان است.

H_1 : " " " " " متفاوت است.

CamScanner

2 b we have more than two variable, and data has drawn independently from Distribution, Condition checked for ANOVA Test.

H₀ : $\mu_A = \mu_B = \mu_c$ ⇒ 3 diffrent diet affect weight equally

H_A : at least one of Diets is diffrent. level of significance ↳ %5

Anova Table:

	Df	SumSq	Mean Sq	F_{value}	P_{value}
Class	2	71.1	35.55	6.197	0.003
Residuals	75	430.2	5.74		

	mean	n
Diet 1	3.3	24
Diet 2	3.02	27
Diet 3	5.15	27

$$SSG = \sum n_j (\bar{y}_j - \bar{y}) = 71.1$$

$$SST - SSG = SSE = 430.2$$

$$\frac{SSG}{Df_{Class}} = Mean\ Sq_{Class} = 35.55$$

71.1, 2

$$\frac{SSE}{Df_{Res}} = Mean\ Sq_{Res} = 5.74$$

430, 75

$$F = \frac{35.55}{5.74} = 6.197 \Rightarrow P_{value} = 0.003$$

Null hypothesis is Rejected

at least there is one Diet which is significantly diffrent.

2.e

I write test for one pair others are same.

H_0 : $\mu_A = \mu_B$ → two group doesn't differ signifantly and two diets has performed same

$\mu_A \neq \mu_B$ → one diet has performed signifiantly Diffrent

(Rest of 2 are same)

$\bar{x}_1$ & $\bar{x}_2$: $t = \dfrac{\bar{x}_1 - \bar{x}_2}{SE} = 0.411$ → Null hypothesis Can not be rejected

$\bar{x}_1, \bar{x}_3$: $t = \dfrac{\bar{x}_1 - \bar{x}_3}{SE} = 3.316$ → two side $0.004 < 0.05$ ✓

Null Reject

$\bar{x}_2, \bar{x}_3$: $t = \dfrac{\bar{x}_2 - \bar{x}_3}{SE} = -2.8462$ → two side $0.02 < 0.5$ ✓

Null Rejected

CamScanner

a $n = 100$, $X = 38$; $\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$

3

Expected	Actuall
50	38
50	62

$$\chi^2 = \frac{(50-38)^2}{50} + \frac{(50-62)^2}{50} = 5.76$$

$\xrightarrow{\text{using R}}$ pchisq (5.76, 1 , lower.tail = FALSE) = $\boxed{0.016}$ → p-value

b distribution of test statistic under Null hypothesis is chi^2 with one degree of freedom

c I have calculated p value using R, but using table we can see in first row, 5.76 is between 97.5% and 99% so our p-value using table is about 2%

d

Expected	Actuall
$\frac{n}{2}$	x
$\frac{n}{2}$	$n-x$

$$\chi^2 = \frac{(x - \frac{n}{2})^2}{\frac{n}{2}} + \frac{(n - x - \frac{n}{2})^2}{\frac{n}{2}}$$

$$= \frac{(x - n/2)^2}{n/2} + \frac{(n/2 - x)^2}{n/2}$$

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

As we can see it is increasing function of $|x - \frac{n}{2}|$

4

Sum Total Player : 359

a H_0 : Nothing is going on! there is no relation between month that some one is born and bieng a pro football player.

H_1 . Iranian people who born early in the year are more likly to be successfull.

b 1. Independence : Samples are independent; ✓

1 randomly sample

2 $n < 10\%$ population

3 each case only contributes to one cell

2. Sample Size : Eeach particular cell has at least 5 expected case. ✓

c

Actuall	147	110	52	50
Expected	85	93	93	88

$$\chi^2 = \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i} = \frac{(147-85)^2}{85} + \frac{(110-93)^2}{93} + \frac{(52-93)^2}{93} + \frac{(50-88)^2}{88} = 80.2348$$

p - value $\simeq 0$

d I got same result in R → P-value $= 2.73 \times 10^{-17}$

e Since we got extremely small p-value, we reject Null hypothesis with high certainty. So, there is a meaning full relation between birth month and being pro football player. Here p-value means small prob that players having that birth date given there is no relation between birthdate and being pro player

5

Null hypothesis: IQ of EE and CE students are equall

Alternative " : " " " " " are not equall

Major	IQ	Rank
EE	125	9
EE	127	13.5
EE	126	11
EE	126	11
EE	105	7
EE	128	15
EE	127	13.5
EE	126	11
EE	79	5
EE	124	8
CE	131	17
CE	129	16
CE	77	4
CE	52	1
CE	134	20
CE	135	19
CE	94	6
CE	68	3
CE	67	2
CE	132	18

$n_{EE} = n_{CE} = 10$

$T_{EE} = 104$; $T_{CE} = 106$

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - T_1$$

$$\left.\begin{array}{l} U_1 = 51 \\ U_2 = 49 \end{array}\right\} \quad U = \min(U_1, U_2) = 49$$

$$\sigma_u = \sqrt{\frac{n^2(2n+1)}{12}} = \cancel{9.57}\ 13.22$$

P value = 0.96

چون بی‌جوری برعکس داره انجام میشه!
می‌توان دید که شرایط خیلی خوبی برای
P value داریم و می‌توانیم شک کنیم
که IQ EE و CE یکسان است.

6 a $X_1, \dots, X_2 \sim \mathcal{N}(\mu, 6^2)$ $\quad\theta = \mu, 6)$

$P(x|\theta) \rightarrow$ liklihood $\Rightarrow P(x_1 \dots x_n|\theta)$

maximize $P(x_1 \dots x_n | \mu, 6) \overset{iid}{=} P(x_1|\mu, 6) \dots P(x_n|\mu, 6)$

$= \prod_{i=1}^{n} P(x_i | \mu, 6)$

$\xrightarrow{\text{log liklihood}} \sum_{i=1}^{N} \log P(x_i|\mu, 6)$

$= -n \log \sqrt{2\pi}\, 6 + \sum_{i=1}^{n} \log e^{\frac{(x_i-\mu)^2}{26^2}} = -n \log \sqrt{2\pi}\, 6 + \frac{\sum (x_i-\mu)^2}{26^2}$

$\Rightarrow \ell(x | \mu, 6) = -n \log\sqrt{2\pi} - n \log 6 + \sum \frac{(x_i-\mu)^2}{26^2}$

$\frac{\partial}{\partial\mu} = 0 \Rightarrow -2 \sum (x_i - \mu) = 0 \Rightarrow +2n\mu = 2\sum x_i$

$\Rightarrow \mu = \frac{\sum x_i}{n}$

$\frac{\partial}{\partial 6} = 0 \Rightarrow \frac{-n}{6} + \frac{\sum (x_i-\mu)^2}{2} \frac{-\cancel{3}2}{6^3} = 0$

$\Rightarrow -n6^2 + \frac{\sum(x_i-\mu)^2}{\cancel{2}} \cancel{2} = 0$

$6^2 = \frac{\sum(x_i-\mu)^2}{n}$

$6_{ME} = \sqrt{\frac{1}{n} \sum (x_i - \mu)^2}$

6.b) S is a random variable with $n-1$ DF since we have n independent samples, and one should be subtracted because we are estimating one parameter which is mean. in other words $n-1$ samples can vary and the nth sample is fixed because it can be obtaind from μ.

6.c : first we calculate mean $\rightarrow \bar{x} = \frac{\sum x_i}{n}$

now estimated SE : $S^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$

$$t_{stat} = \frac{x - \bar{x}}{S}$$

after obtaining t ; we will find it's value from t student distribution (with n-1 degree of freedom) then we will compare it with our significance level α

7.a : since we have independent samples we use wilcoxon rank sum test

	time	ranks
gp1	19	13
	14.4	7
	18.2	11
	15.6	9
	14.5	8
	11.2	2
	13.9	6
	11.6	3.5
gp2	12.1	5
	19.1	14
	11.6	3.5
	21.0	16
	16.7	10
	10.1	1
	18.3	12
	20.5	15

$T_1 = 59.5$

$U_1 = n^2 + \frac{n(n+1)}{2} - T_1 = 40.5$

$\boxed{\min(U_1, U_2) = 23.5}$

$T_2 = 76.5$

$U_2 = n^2 + \frac{n(n+1)}{2} - T_2 = 23.5$

$\mu_u = \frac{n^2}{2} = 32$; $\sigma_u = \sqrt{\frac{n^2(2n+1)}{12}} = 9.522$

5.

$U = 23.5$; $n = 8 \Rightarrow 23.5 > 13$

we can not Reject Null Hypothesi

P value $= 0.4$

b

$W_A = 59.5$

$W_B = 76.5$

$U_A = 59.5 - \frac{8 \times 9}{2} = 59.5 - 36 = 23.5$

$U_B = 76.5 - 36 = 40.5$

$\left.\begin{array}{l} \min(U_A, U_B) = 23.5 \\ \text{willcox}(8,8) = 13 \end{array}\right\}$ we cannot Reject

c

for smoothly combining two sets we can fit a Gaussian Kernel for each data point the we can see our distribution has less uncountinuity. though we have to decide about Kernel size, for smoother approximation we use bigger kernel size and if we have relativly lots of data points we can use smaller Kernel size.

6

b Since we have two dependent samples, and we want to know whether show was effective or not; Wilcoxon Signed Rank Test is used.

Before	After	Diff	Rank
60	58	2	1.5
56	58	-2	-1.5
80	83	-3	-3.5
73	67	6	6
14	17	-3	-3.5
32	36	-4	-5

$W^+ = 7.5$

$W^- = 7+5+1.5 = 13.5$ $\Rightarrow$ W value $= 7.5$

for significance level 0.05 $\longrightarrow$ one tail: $\alpha = 2$
two tail: $\alpha = 0$

Null hypothesis: Central tendency of their opinion has not changed.

Alternative hypothesis: Central " " " has changed.

Note: In question has been asked "determine is there a significant increase in number of supporters" so one tail test should be used, But the data shows that talk may reduce his supporters so two tail test is been used

8 continued

$W_{stat} > \alpha|_{0.05} \longrightarrow$ We dont reject Null hypothesis

we cannot approximate with normal since number of observation is not enough ; $p \simeq 0.6$

b implementation in R has been attached with files & I got same Result

c

Befor	After	Diff	R
78	65	13	7
50	20	30	8
40	50	-10	-4.5
49	8	32	9
20	16	4	1
50	44	6	3
50	38	12	6
50	40	10	4.5
50	45	5	2

$W^T = 40.5$

$W^- = 4.5$ $\Rightarrow W_{STat} = 4.5$

Null Hypothesis: Central tendency of thier opinion has not changed.

Alternative " : " " " " " has changed.

for $\alpha = 0.05$ and two tail test $W|_{N=9} = 5$

$4.5 < 5 \longrightarrow$ Result is significant.

$0.03 = P < 0.05$

9

8 d No we can not Relate the results since they are two independent tests;

10 a Chi^2 test for independenc is good in order to show relation between two groups. Since we have two categorical groups and we want to know the relation

10 b No. In order to find causal relation-ship we need to notice two steps:

1 random Sampling

2 Experiment setup

if we sample from hospital, there is higher chance to Sample people with bone disease or respiratory since we are in hospital! therefore we can not take conclusion from Hospitalized people that bone disease has correlation with Respiratory disease. General population is better to experiment.