نام درس:ناپارامتری

استاد راهنما:استاد دکتر حسینی

گرداورنده:رضا باقری فهرجی

شماره دانشجویی: 9611350005

موضوع پروژه:ازمون های ناپارامتری با نرم افزار R

سوال اول: فرض کنید نمونه ای 30 تایی از توزیع پیوسته و اکیدا صعودی استخراج شده است.

0.11 0.14 0.16 0.19 0.26 0.28 0.33 0.38 0.38

0.52 0.58 0.62 0.63 0.76 0.86 0.87 0.88 0.91

0.92 0.94 0.95 1.01 1.15 1.15 1.19 1.21 1.46

1.73 1.99

الف ـ دهک هفتم توزیع را با استفاده از نمونه فوق براورد کنید

ب ـ یک فاصله اطمینان95درصد برای میانه توزیع بیابید

حل الف-

```
> x<-c(0.11,0.14,0.16,0.19,0.26,0.28,0.33,0.38,0.38,0.52,0.58,0.62,0.63,0.76,
0.86,0.87,0.88,0.91,0.92,0.94,0.95,1.01,1.15,1.15,1.19,1.21,1.46,1.72,1.73,1.
99)
> x<-sort(x)
> p<-0.7
> n=length(x)
> r<-floor((n+1)*p)
> w<-(n+1)*p-r
> Qp<-(1-w)*x[r]+w*x[r+1]
> Qp
[1] 0.992
```

حل ب –.

```
> a<-0.95
> p<-0.5
> i<-{}
> M<-matrix(0,n*(n-1)/2,3)
> for(j in 1:(n-1)){
+   for (k in (j+1):n) {
+
+
+   i=c(i,k)
+   l=length(i)
+ M[l,1]=j
+ M[l,2]=k
+ M[l,3]=pbinom(k-1,n,p)-pbinom(j-1,n,p)
+ }
+ }
> M
```

با ران کردن M ماتریسی با 333سطر میدهد که به دلیل اینکه فضای زیادی را در بر میگیرد از این قسمت حذفشان کردیم و فقط نتیجه فاصله اطمینان را مینویسیم.

```
> p=M[,3]
> a1=min(p[p>=a])
> c=M[M[,3]==a1,]
> dimnames(c)=list(NULL,c("j","k","1-alpha"))
```

```
> c
     j  k  1-alpha
[1,]  6 20 0.950469
[2,] 11 25 0.950469
```

تحلیل:در قسمت الف که دهک مربوطه توزیع را بدست اوردیم و در قسمت ب فاصله اطمینان را حساب کردیم و توضیح مختصری در مورد فاصله اطمینان در زیر میدهیم.

فاصله اطمینان: یک فاصله یا محدوده‌ای عددی، شامل پارامتر مورد جامعه است. پس می‌توان فاصله اطمینان را نوعی برآورد فاصله‌ای در نظر گرفت.

ازمون علامت(نشانه): SIGN-Test

در آزمون نشانه می خواهیم برای توزیع پیوسته و صعودی F(x) آزمون آماری زیر را در مرود چندک مرتبه P ام انجام دهیم.

تعریف چندک p جامعه:

$p=p(x<Q_p)$

مرحله یک: فرضیه ازمون

$$\begin{cases} HO:QP=a \\ H1:QP>a \end{cases}$$

$$\begin{cases} HO:QP=a \\ H1:QP<a \end{cases}$$

$$\begin{cases} HO:QP=a \\ H1:QP\neq a \end{cases}$$

مرحله دوم:سطح ازمون را α در نظر میگیریم.یعنی

P(RH0|H0 درست)=α

مرحله سوم:نمونه تصادفی X1 تا Xn از توزیع F_X x را اختیار میکنیم.

مرحله چهارم:اماره ازمون را تعریف میکنیم.

مرحله پنجم:تعریف ناحیه بحرانی

1 ـ اگر فرض مقابل QP=a ، H1:QP>=a باشد انگاه:

If $B\geq K_\alpha$ RHO

2- اگر فرض مقابل QP=a ، H1:QP<a باشد انگاه:

If $B \leq K_{\alpha}$ RH0

3- اگر فرض مقابل QP=a ، H1:QP≠a باشد انگاه:

If $B \geq K_{\alpha 1}$ OR $B \leq K_{\alpha 2}$ RH0

سوال دوم: فرض کنید نمونه ای 30 تایی از توزیع پیوسته و اکیدا صعودی استخراج شده است.

0.11 0.14 0.16 0.19 0.26 0.28 0.33 0.38 0.38

0.52 0.58 0.62 0.63 0.76 0.86 0.87 0.88 0.91

0.92 0.94 0.95 1.01 1.15 1.15 1.19 1.21 1.46

1.73 1.99

ازمون
$$\begin{cases} H0: Q_{0.5} = 1 \\ H1: Q_{0.5} < 1 \end{cases}$$
را در سطح α=0.05 ازمون کنید.(ازمون علامت نشانه)

```
> x<-c(0.11,0.14,0.16,0.19,0.26,0.28,0.33,0.38,0.38,0.52,0.58,0.62,0.63,0.76,
0.86,0.87,0.88,0.91,0.92,0.94,0.95,1.01,1.15,1.15,1.19,1.21,1.46,1.72,1.73,1.
99)
> library(BSDA)

> SIGN.test(x,md=1,alternative = "less")

        One-sample Sign-Test

data:  x
s = 9, p-value = 0.02139
alternative hypothesis: true median is less than 1
95 percent confidence interval:
      -Inf 0.9397518
sample estimates:
median of x
      0.865

Achieved and Interpolated Confidence Intervals:

                    Conf.Level  L.E.pt  U.E.pt
Lower Achieved CI       0.8998    -Inf  0.9200
Interpolated CI         0.9500    -Inf  0.9398
Upper Achieved CI       0.9506    -Inf  0.9400
```

تحلیل: روش کار به این نحوه بود که در بین 30 عدد بالا با استفاده از ازمون علامت (نشانه) عدد های بزرگتر مساوی یک را محاسبه کرده و انها را به عنوان S در نظر گرفته شده که تعداشون 9 عدد است.و برای محاسبه P_value حالت دوم که در بالا نوضیح دادیم یعنی $B \leq K_{\alpha}$ رخ میدهد که برابر 0.02139 می شود.

سوال سوم:داده های زیر اخرین مقاومت بر حسب کیلوگرم برای یک نوع مفتعول است که در یک ازمایشگاه تهییه شده اند.

163 165 162 189 161 171 158 151 169 182 163 139 172 165 148 166 172 163 187 173

ازمون زیر را انجام دهید.(ازمون علامت نشانه)

$$\begin{cases} H0:Q_{0.5}=160 \\ H1:Q_{0.5}>160 \end{cases}$$

```
> y<-c(163,165,162,189,161,171,158,151,169,182,163,139,172,165,148,166,172,16
3,187,173)
> library(BSDA)
> SIGN.test(y,md=160,alternative = "greater")

        One-sample Sign-Test

data:  y
s = 16, p-value = 0.005909
alternative hypothesis: true median is greater than 160
95 percent confidence interval:
 162.7928      Inf
sample estimates:
median of x
        165

Achieved and Interpolated Confidence Intervals:

                  Conf.Level   L.E.pt U.E.pt
Lower Achieved CI     0.9423 163.0000    Inf
Interpolated CI       0.9500 162.7928    Inf
Upper Achieved CI     0.9793 162.0000    Inf

> p=0.5
> A=y[y!=160]
> n=length(A)
> B=length(A[A>160])
> B
[1] 16
> p_value=pbinom(b,n,1-p)
> p_value
[1] 0.9987116
```

تحلیل: این سوال هم طبق ازمون علامت نشانه حل شده و در اینجا عدد مقایسه ما با مشاهدات 160 می باشد که در بین 20 عدد، 16 عدد از مشاهدات ما بزرگتر است که B=16 و برای محاسبه P_value از حالت اول میرویم که $B \geq K_\alpha$ می باشد که حاصل ان برابر 0.005909 می شود

سوال چهارم:برای داده های زیر ازمون را برای $\alpha=0.1$ را انجام دهید.

$$\begin{cases} H0:Q_{0.75}=85 \\ H1:Q_{0.75}\neq 85 \end{cases}$$

89 90 86 80 97 81 94 82 87 93 94 84 83 78 98

```
> z<-c(89,90,86,80,97,81,94,82,87,93,94,84,83,78,98)
```

```
> p=0.75
> A=z[z!=85]
> n=length(A)
> b=length(A[A>85])
> d<-pbinom(b,n,1-p)
> a<-1-pbinom(b-1,n,1-p)
> p_value<-2*min(a,d)
> p_value
[1] 0.008386029
> B=length(z[z>85])
> B
[1] 9
```

تحلیل: این سوال هم از ازمون علامت(نشانه)حل شده فقط در مقایسه با دو مثال قبل این تفاوت را دارا است که ازمون ما دو طرفه می باشد یعنی p_value ضرب در 2 می شود یعنی p_value ما 0.004 بود که وقتی در 2 ضرب شده برابر 0.008 شده است .

ازمون رتبه ای:(ازمون رتبه –ویلکاکسون)

روش حل:

1_قرار میدهیم $D_i=X_i-V$

2_|Di| ها را مرتب کرده و رتبه انها را بدست می اوریم و فرض میکنیم که Ri رتبه Di باشد و قرار می دهیم.

$$Z_i=\begin{cases}1 & D_i>0\\ 0 & D_i<0\end{cases}$$

3_امار مورد نظر به صورت زیر تعریف می شود.

$T=W=\Sigma R_i Z_i$

سوال پنجم: مقادیر مشاهده شده یک نمونه تصادفی ده تایی از توریعی که نسبت به v متقارن است عبارتند از:

فرضیه زیر را در سطح α=0.05 ازمون کنید.

$$\begin{cases}H_0:V=8\\ H_1:V>8\end{cases}$$

10.2 14.1 9.2 11.3 7.2 9.8 6.5 11.8 8.7 10.8

```
> w<-c(10.2,14.1,9.2,11.3,7.2,9.8,6.5,11.8,8.7,10.8)
> y<-8
> wilcox.test(w-y,alternative = "greater")

        Wilcoxon signed rank exact test

data:  w - y
```

```
V = 49, p-value = 0.01367
alternative hypothesis: true location is greater than 0

> R<-1-psignrank(48,10)
> R
[1] 0.01367188
```

با استفاده از دستور ()Wilcox.test ازمون فوق را انجام میدهیم م مقدار w=49 و p_value=0.01367 می شود.

ازمون جمعی-رتبه ای – ویلکاکسون:

آزمون جمعی رتبه ای ویلکاکسون (که معادل آزمون من – ویتنی میباشد)برای دو جامعه مستقل از داده ها کاربرد دارند.

X~F(x)　　　　　Y~G(y)

و فرضیه ان به صورت زیر است:

$$\begin{cases} H0:C=0 \\ H1:C\neq 0 \end{cases} \quad \begin{cases} H0:C=0 \\ H1:C<0 \end{cases} \quad \begin{cases} H0:C=0 \\ H1:C>0 \end{cases}$$

ابتدا فرضیه زیر زیر را در نظر می گیریم و سپس برای ازمون کردن فرضیه در α مراحل زیر را انجام میدهیم:

مرحله اول: جامعه ها را با هم مخلوط میکنیم نمونه ای به حجم N= m+n از x ها و y ها داریم

مرحله دوم:نمونه های مخلوط شده با هم را مرتب می کنیم و رتبه Xi را Ri و Yi را Sj می نامیم.

مرحله سوم: Wx را بست می اوریم:

$$Wx=\Sigma Ri$$

مرحله چهارم:فرض H0 در صورتی رد خواهد شد که که H1 درست باشد.

سوال ششم:از هشت نفر خواسته شده در حالت عادی و استرس زا یک پازل ساده را بچینند.در حالت استرس زا به انها گفته شده است که 3 دیقه به انها شوک خفیف پس از اغاز ازمایش و 30ثانیه پس از ان پازل به انها وارد خواهد شدو زمانی که به پایان رسید فشار خون انها اندازگیری می شود .داده ها در جدول زیر ارائه شده است.ایا شواهد کافی در اطلاعات موجود است که نشان دهد تحت شرایط استرس زا فشار خون بیشتر است و ازمون رتبه علامت ویلکاکسون استفاده کنید.

A normal:126 117 115 118 118 128 125 120

B esteres:130 118 125 120 121 125 130 120

```
> x<-c(126,117,115,118,118,128,125,120)
> y<-c(130,118,125,120,121,125,130,120)
> sort(x)
[1] 115 117 118 118 120 125 126 128
> sort(y)
[1] 118 120 120 121 125 125 130 130
> wilcox.test(y,x,correct = T,paired = T,alternative = "greater")

        Wilcoxon signed rank test with continuity correction

data:  y and x
```

```
V = 24.5, p-value = 0.0452
alternative hypothesis: true location shift is greater than 0
```

تحلیل:ابتدا دو نمونه مورد نظر را وارد کردیم و سپس انها را sortکردیم و با استفاده از دستور wilcox.test ان را ازمون میکنیم دستور correct=T یعنی تصیح پیوستگی انجام شود. اگر تصیح پیوستگی نخواهیم انجام شود از F استفاده میکنیم. وستور paired=T یعنی ازمون جفتی انجام میشود که باید مقدار دو نمونه برابر باشد.در غیر این صورت از paired=Fاستفاده میکنیم.

سوال هفتم:از 10 بیمارستان داده های بالینی مربوط به اثر بخشی دو دارو برای درمان یک نوع بیماری جمع اوری شده است.ایا اطلاعات تفاوت در اختلاف در میزان بازیابی دو گروه را نشان میدهد.(ازمون علامت)

B	A	بیمارستان
85.4	75	1
83.1	69.5	2
80.2	85.7	3
74.5	74	4
70	69	5
81.5	83.3	6
75.4	68.9	7
78	77.8	8
85.4	72.2	9
80.4	77.4	10

```
> x<-c(75,69.8,85.7,74,69,83.3,68.9,77.8,72.2,77.4)
> y<-c(85.4,83.1,80.2,74.5,70,81.5,75.4,78,85.4,80.4)
> d=y-x
> d=d[d=!0]
> n=length(d)
> B=length(d[d>0])
> p_value=2*min(pbinom(B,n,1/2),1-pbinom(B-1,n,1/2))
> p_value
[1] 0.109375
```

تحلیل:داده ها فوق را با نرم افزار به روش ازمون علامت انجام دادایم و مشاهدات فوق را بدست اوردیم که در ان P_value=0.109375 را بدست اوردیم.

سوال هشتم:امتیاز های دو گروه A و B از دانشجویان که به ترتیب 10 و 10 می باشند .در نظر بگیرید.نتایج یک ازمون بر ای این دانشجویانن به صورت امتیاز های عالی ، متوسط و ضعیف می باشد.نتایج ازمون به تفکیک گروه در جدول زیر امد ه است. در سطح
معنی داری 0.05 با استفاده از ازمون جمعی رتبه ای ویلکاکسون بررسی کنید.ایا دو گروه بر اساس این ازمون یکسانند .

$$\begin{cases} H0:C=0 \\ H1:C\neq 0 \end{cases}$$

:

	عالی	متوسط	ضعیف
A	4	3	3
B	3	5	2

```
> A<-c(rep(1,4),rep(2,3),rep(3,3))
> B<-c(rep(1,3),rep(2,5),rep(3,2))
> wilcox.test(A,B,alternative = "two.sided",correct = T)

        Wilcoxon rank sum test with continuity correction

data:  A and B
W = 49.5, p-value = 1
alternative hypothesis: true location shift is not equal to 0
```

تحلیل: در بالا دو گروه A و B را در سه حالت مختلف عالی،متوسط و ضعیف ازمون کردیم با ازمون جمعی رتبه ای ویلکاکسون که در بالا هم شرح داده بودیم را ازمون کردیم با استفاده از دستور wilcox.test و در اینجا از two.sided استفاده کردیم چون ازمونی که میخواستیم انجام دهیم یک ازمون دو طرفه بود . دستور correct=T که قبلا هم در مثال قبل اشاره شد برای اینکه تصیح پیوستگی انجام شود استفاده میشود.که برای مقدار P_value=1 چون ازمون دو طرفه بود ضرب در عدد 2 شده است یعنی P_value=0.5 بوده است.

ازمون من – ویتنی:

ازمون دیگری که می توان بر اساس ان دو جامعه مستقل را مقایسه کرد ازمون من – ویتنی است که به ازمون u معروف است.

اگر F(x) و G(y) به ترتیب توابع دو جامعه مذکور باشند و x1,…xm و y1…yn دو نمونه تصادفی باشند. به طوری که F و G پیوسته و نمونه مستقل اند.فرضیلت به صورت زیر است:

$$\begin{cases} H0:F(z)=G(z) \\ H1:F(z)>G(z) \end{cases}$$

برای ازمون فرضیات بالا در سطحα مراحل زیر را انجام میدهیم

مرحله اول:ابتدا دو نمونه را با هم ادغام کرده و سپس مرتب میکنیم

مرحله دوم:متغیر تصادفی زیر را در نظر می گیریم

$$Dij=\begin{cases} 1 \quad \text{if} \quad yi<xi \quad \text{سمت چپ x y} \quad & i=1\ldots m \\ 0 \quad \text{if} \quad yj>xj \quad \text{سمت راست x y} \quad & j=1\ldots n \end{cases}$$

اماره ازمون به صورت زیر است:

$u=w_{xy}=\Sigma\Sigma D_{ij}$

جمع تعداد xهایی که در راست y هستند.

$u=w_{yx}=\Sigma\Sigma D_{ij}$

جمع تعداد y هایی که در راست x هستند.

سوال نهم:پانزده باتری تجربی به طور تصادفی از کارخانه A و پانزده باتری استاندارد به طور تصادفی از کارخانه B انتخاب شده است و به طور همزمان همه این 30 باتری تحت بارالکتریکی یکسان قرار داده شده اولینباتری خراب شده از کارخانهد A ،دومی از کارخانه B، سومی از کارخانه B ...

دنباله زیر:

A B B B A B A A B B B B A B A B B B B A A B A A A B A A A A

با استفاده از ازمون من – ویتنی تعیین کنید ایا شواهد کافی برای پذیرش این نتیجه که طول عمر تجربی A بیشتر از باتری استاندارد استوجود دارد؟ 0.05=α

HO:A طول عمر باتری =Bطول عمر

H1:Aطول عمر>Bطول عمر

```
> A<-c(1,5,7,8,13,15,20,21,23,24,25,27,28,29,30)
> B<-c(2,3,4,6,9,10,11,12,14,16,17,18,19,22,26)
> n<-length(A)
> m<-length(B)
> N<- n+m
> C<-c(B,A)
> R=rank(C)
> WB=sum(R[1:n])
> WB
[1] 189
> WAB=WB-n*(n+1)/2
> WAB
[1] 69
> EWB=n*(N+1)/2
> VWB=n*m*(N+1)/12
> p_value=pnorm(WB+1/2-EWB)/sqrt(VWB)
> p_value=pwilcox(WAB,n,m)
> p_value
[1] 0.03710225
```

تحلیل: برای محاسبه با ازمون من – وتینی کافی است انها را ترکیب کرده و به ترتیب به هر یک انها رتبه میدهیم و رتبه های مربوط به A را جدا و رتبه های مربوط به B را جدا در نرم افزار تعریف میکنیم و برای هر یک روند فوق را ادامه میدهیم WBهمان جمع رتبه های B می باشد و WAB به روش زیر بدست می اید.

$WAB=WB-\frac{n(n+1)}{2}$

آزمون کراسکال والیس(kruskal.test):

آزمون کراسکال والیس برای انجام مقایسه در مورد بیش از دو جامعه مستقل مورد استفاده قرار می گیرد در این آزمون می خواهیم بیازماییم متغیرهای مستقل$x_1,x_2,...,x_n$هم توزیع هستند یا نه؟

یک نمونهni تایی از هر Xiدر نظر می گیریم این نمونه هارا ترکیب میکنیم تا یک نمونه $n=n_1+n_2+...+n_k$تایی به دست آید و

R:معدل تمام رتبه ها

R_i:معدل رتبه ها برای نمونه iام

$H=12/n(n+1)\Sigma Ri^{2}/ni-3(n+1)$

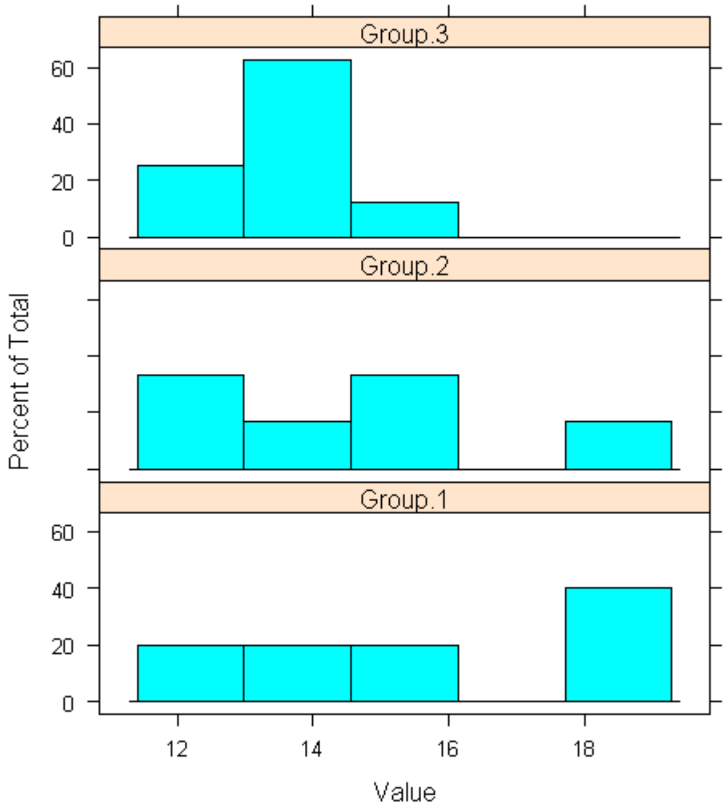
سوال دهم:در جدوا زیر نمرات دانشجویان ورودی ازمون دکترای سه دانشگاه A و B و C داده شده است می خواهیم با $\alpha=0.05$ بیازمایم

A: 19 11.7 17.8 14.8 13.9

B: 18.2 14.8 13.1 12.6 15.2 12.8

C:13.4 14.1 12.3 12.3 14.7 13.9 13.8 14.3

```
> Input =("
+ Group Value
+ Group.1 19
+ Group.1 11.7
+ Group.1 17.8
+ Group.1 14.8
+ Group.1 13.9
+ Group.2 18.2
+ Group.2 14.8
+ Group.2 13.1
+ Group.2 12.6
+ Group.2 15.2
+ Group.2 12.8
+ Group.3 13.4
+ Group.3 14.1
+ Group.3 12.3
+ Group.3 12.3
+ Group.3 14.7
+ Group.3 13.9
+ Group.3 13.8
+ Group.3 14.3
+ ")
> Data = read.table(textConnection(Input),header=TRUE)
> library(dplyr)
> Data =mutate(Data,Group = factor(Group, levels=unique(Group)))
> library(lattice)
> histogram(~Value|Group,data=Data,layout=c(1,3))
```



تحلیل:ابتدا سه گروه را با اعداد 1و2و3 تعیین کردیم سپس و با دستور READ.TABLE به صورت جدول بندی شده فرخ وانی کردو سپس نمودار هیستوگرام سه گروه رسم شد که می توان ان را از نمودار هیستوگرام ان بررسی های مختلفی را ا نجام داد.

```
 > library(FSA)
> Summarize(Value ~ Group,data = Data)
    Group n  mean        sd  min     Q1 median    Q3  max
1 Group.1 5 15.44 2.9585469 11.7 13.900  14.80 17.80 19.0
2 Group.2 6 14.45 2.1314314 12.6 12.875  13.95 15.10 18.2
3 Group.3 8 13.60 0.8864053 12.3 13.125  13.85 14.15 14.7
> kruskal.test(Value ~ Group,data = Data)

        Kruskal-Wallis rank sum test

data:  Value by Group
Kruskal-Wallis chi-squared = 1.663, df = 2, p-value = 0.4354
```

تحلیل:در این قسمت نیز از دستور summary استفاده شد که اطلاعاتی از جمله میانگین،میانه،چارک اول و سوم و انحراف معیار را بدست اورده شده است که می توان اطلاعات مهمی را از ان کسب کرد.و همچنین مقدار P_value بدست امده است.

ازمون فریدمن (friedman.test):

یکی از آزمون‌های آماری است که برای مقایسه چند گروه کاربرد دارد و از نظر میانگین رتبه‌های گروه‌ها را معلوم می‌کند که آیا این گروه‌ها می‌توانند از یک جامعه باشند یا نه؟

آزمون فریدمن یک آزمون ناپارامتری، معادل آنالیز واریانس با اندازه‌های تکراری (درون گروهی است) که از آن برای مقایسه میانگین رتبه‌ها در بین k متغیر (گروه) استفاده می‌کنیم. فرض کنید می خواهیم از یک نمونه شامل 10 نفر در مورد 5 کالا نظر خواهی کنیم-یعنی از آن‌ها بخواهیم که به هر یک از کالاها از نظر کیفیت امتیاز بدهند-سپس میانگین امتیازات کالاها را باهم مقایسه و بررسی کنیم که اگر اختلافات امتیازات کالاها معنی دار است کدام کالا بیشترین امتیاز و کدام کالا کمترین امتیاز را کسب کرده‌است. در چنین حالتی شما با 5 متغیر روبرو هستید که این متغیرها از لحاظ آماری به هم وابسته هستند. زیرا اندازه‌هایی هستند که توسط هر نمونه تکرار شده‌اند.تفاوت آنالیز واریانس با اندازه‌های تکراری (درون گروهی) با آزمون فریدمن در این است که در آنالیز واریانس شما از هر نمونه یک متغیر را به صورت تکراری در حالات مختلف اندازه‌گیری می‌گیرید. (مثلا از هر نفر (نمونه) در سه حالت ایستاده، نشسته و دراز کش فشار خون را اندازه‌گیری می‌کنید). در صورتی که در آزمون فریدمن هر یک از نمونه‌ها امتیازی را به چند گروه (شی یا فرد یا..) اختصاص می‌دهند. در هر دوی این آزمون‌ها متغیرها، توسط نمونه‌ها مقدار گرفته‌اند ولی نکته مورد اختلاف این است که در آنالیز واریانس، در یک نمونه اندازه‌ها تکراری هستند ولی در آزمون فریدمن اندازه‌ها، امتیازات داده شده توسط یک نمونه است. در آزمون فریدمن فرض H_0 مبتنی بر یکسان بودن میانگین رتبه‌ها در بین گروه هاست. رد شدن فرض صفر به این معنی است که در بین گروه‌ها حداقل دو گروه با هم اختلاف معنا داری دارند.

سوال یازدهم:چهار استاد،سه دانشجو را برای ازمون ورودی دکترا مصاحبه عملی کردندبا ازمون فریدمن ازمون کنید مصاحبه کنید نحوه محاسبه یکسان بوده است.

دانشجو / استاد	1	2	3	جمع سطری
1	3	4	3	10
2	4	3	4	11
3	2	2	1	5
4	1	1	2	4
	10	10	10	30

```
> x1<-c(3,4,3)
> x2<-c(4,3,4)
> x3<-c(2,2,1)
> x4<-c(1,1,2)
> x<-cbind(x1,x2,x3,x4)
> friedman.test(x)

        Friedman rank sum test

data:  x
Friedman chi-squared = 7.4, df = 3, p-value = 0.06018
```

تحلیل:هر یک از نظر استاد ها را در یک متغیر ذخیره کردیم و با استفاده از دستور friedman.test داده ها را ازمون میکنیم.

ازمون برازندگی کولمو اسمیرنف (k-S):

در انتخاب یک آزمون آماری برای تحقیق، باید تصمیم بگیریم که آیا از آزمون‌های پارامتریک استفاده کنیم یا آزمون‌های ناپارامتریک. یکی از اصلی‌ترین ملاک‌ها برای این انتخاب، انجام آزمون کولموگروف-اسمیرنوف است. آزمون کولموگروف-اسمیرنوف،

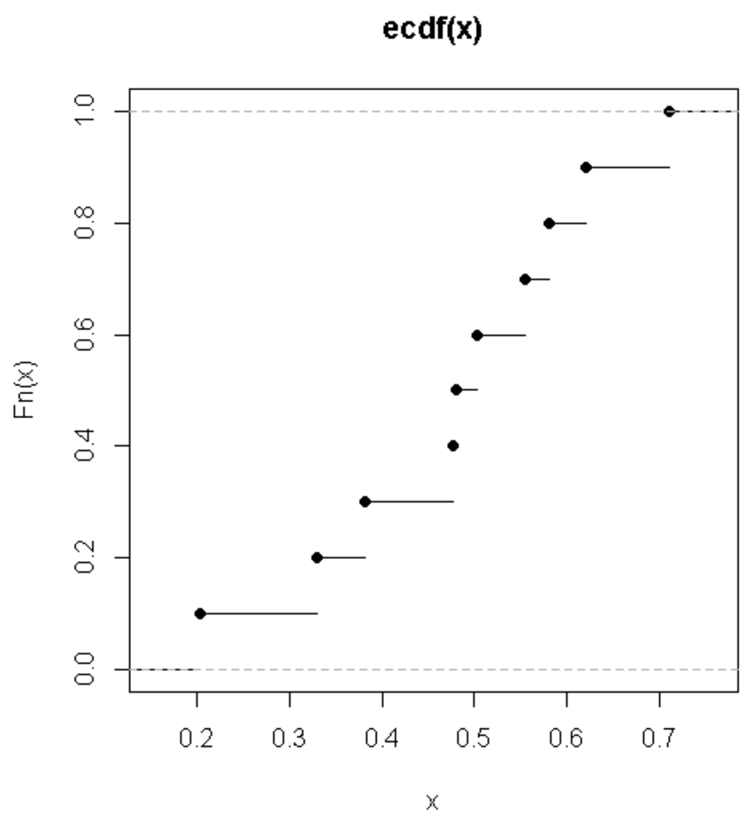
نرمال نبودن توزیع داده‌ها را نشان می‌دهد. یعنی اینکه توزیع یک صفت در یک نمونه را (مثلا سن در بین ۱۰۰ نفر نمونه پرستاران) با توزیعی که برای جامعه، مفروض است (برای مثال سن تمام پرستاران) مقایسه می‌کند. اگر تست کولموگروف-اسمیرنوف رد شود، داده‌ها دارای توزیع نرمال می‌باشند، و امکان استفاده از آزمون‌های آماری پارامتریک برای تحقیق، وجود دارد. بالعکس، اگر تست کولموگروف-اسمیرنوف تأیید شود، یعنی داده‌ها دارای توزیع نرمال نیستند، بنابراین باید از آزمون‌های ناپارمتریک در تحقیق استفاده کنیم.

سوال دوازدهم:نمونه تصادفی 10 تایی به صورت زیر داریم:

0.621 0.503 0.203 0.477 0.710 0.581 0.329 0.480

با میزان α=0.05 بیازمایید که ایا این نمونه از توزیع یکنواخت روی فاصله ی(1و0) ساخته شده است.

```
> x<-c(0.621,0.503,0.203,0.477,0.710,0.581,0.329,0.480,0.554,0.382)
> plot(ecdf(x))
```




```
> ks.test(x, "punif")

        One-sample Kolmogorov-Smirnov test

data:  x
D = 0.29, p-value = 0.3067
alternative hypothesis: two-sided
```

```
> ks.test(x, "pnorm")

        One-sample Kolmogorov-Smirnov test

data:  x
D = 0.58043, p-value = 0.0009996
alternative hypothesis: two-sided
```

تحلیل: داده ها رو وارد کرده و برای اینکه بخواهیم ازمون کولمو اسمیرنف را انجام دهیم از دستور ks.test استفاده میکنیم در اینجا به این دلیل از punif و pnorm استفاده شده است به این دلیل که در صورت سوال اشاره شده است توزیع یکنواخت و همچنین میتوان نمودار پلکانی ان را رسم کرد.

سوال سیزدهم:دریک ازمون ریاضی نمره پنج دختر دیپلمه عبارتنداز:

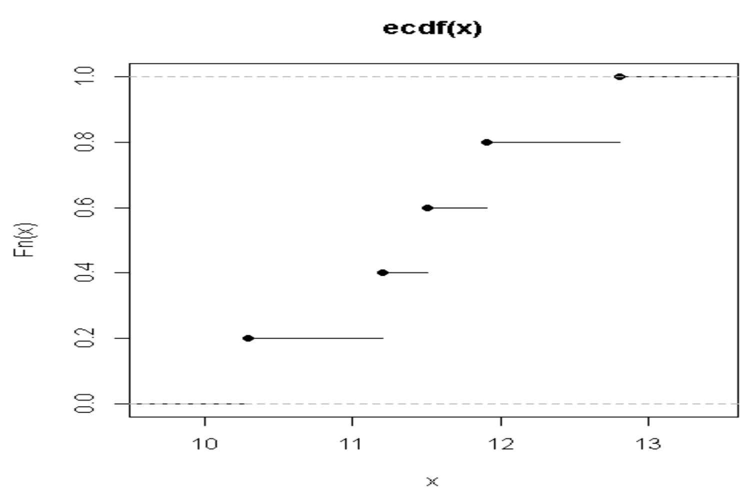
10.3 11.2 11.5 11.9 12.8

و نمره شش پسر دیپلمه عبارتند از:

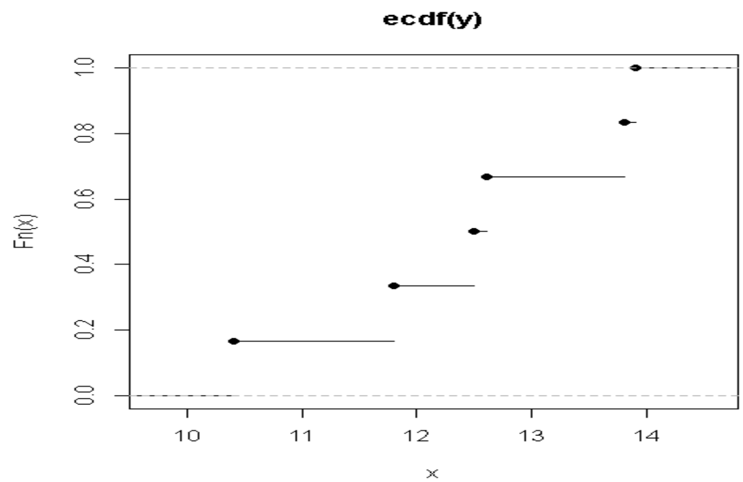
10.4 11.8 12.5 12.6 13.8 13.9

ایا با میزان α=0.05این نمرات هم توزیع هستند؟

```
> x<-c(10.3,11.2,11.5,11.9,12.8)
> y<-c(10.4,11.8,12.5,12.6,13.8,13.9)
> ks.test(x, y, alternative = "two.sided")

        Two-sample Kolmogorov-Smirnov test

data:  x and y
D = 0.46667, p-value = 0.474
alternative hypothesis: two-sided

> plot(ecdf(x))
```




```
> plot(ecdf(y))
```


ecdf(y)
Fn(x)
1.0
0.8
0.6
0.4
0.2
0.0
10
11
12
13
14
x