

Тақырыбы: «Дисперсиялық талдаудың бір факторлы параметрлік емес ұқсастығы критерий- Крускал Уоллис критерийі»

Қабылдаған:

Сейілханова А.Ә

Орындағандар:

2-топша

1. Әбдразақ Әйгерім
2. Ерсінова Арайлым
3. Жұмабай Гүлзира
4. Қасым Айнұр
5. Сыдықова Назым

# *Жоспары:*

## **I. Кіріспе**

## **II. Негізгі бөлім**

- а) Дисперсиялық талдау.
- б) Дисперсиялық талдаудың түрлері.
- в) Крускал-Уоллис критерийі .
- г) Бір факторлы дисперсиялық талдаудың параметрлік емес баламасы. Крускал-Уоллистің критерийі.

## **III. Қорытынды.**

## **IV. Пайдаланылған әдебиеттер.**

## Дисперсия

бұл әртүрлі факторлардың әсерінен туындайтын белгінің өзгергіштігі.

**Дисперсиялық талдау** – бұл екіден артық топтардың орта мәндерін салыстыру үшін, яғни бірнеше тәуелсіз топтардың бір бас жиынтыққа жататындығын немесе жатпайтындығын анықтау үшін қолданылатын талдау әдісі. Орта мәндердің арасындағы айырмашылықтарды анықтау үшін *дисперсиялар* қолданылады.

Дисперсиялық талдауды ағылшын ғалымы, математик және генетик Рональд Фишер 1938 ж. дисперсиялық талдауға мынадай анықтама берген: Дисперсиялық талдау-«бір себептерден болатын дисперсияны екінші бір себептерден болатын *дисперсиядан айыру*».

Дисперсиялық талдаудың мақсаты зерттеліп отырған кездейсоқ шаманың мәніне әсер ететін, шама мәнінің тұрақсыздығын тудыратын факторларды анықтау болып табылады.

**Дисперсиялық  
талдау  
міндеттері:**

белгіленген немесе кездейсоқ болулары мүмкін бірнеше деңгейлермен сипатталатын факторлардың ықпалын зерттеу.

*Факторлық  
белгілер(фактор) –  
зерттелетін  
кұбылысқа ықпал  
ететін белгілер.*

*Нәтижелік белгілер  
(факторға жауап) –  
факторлық  
белгілердің ықпалы  
нәтижесінде өзгертін  
белгілер.*

**Эксперименттік ұйғарымдарда  
дисперсиялық талдаудың  
бірнеше түрлері бар.  
Солардың ішінде жиі кездесетіні мыналар:**

**Бір факторлы**

Бір фактордың ықпалы  
тексерілетін  
дисперсиялық талдау *бір  
факторлы* деп аталады.

**Көп факторлы**

Екі немесе одан да көп  
факторлардың ықпалын  
зерттеу үшін *көп  
факторлы* дисперсиялық  
талдау қолданылады.

## Бір факторлы дисперсиялық талдау

Топтар бір фактордың деңгейлері арқылы анықталады. Бас жиынтықтағы айнымалы әр топта қалыпты таралған және барлық топтардың дисперсиялары бірдей.

Айырмашылықтың шамасын бағалау үшін таңдама орта мәндердің шашылуын топ ішіндегі мәндердің шашылуымен салыстыру қажет.

Салыстырылатын топтар саны фактордың (тәуелсіз айнымалы) деңгейлеріне сәйкес анықталады.

## *Мысалы:*

- фактор - жыл мезгілі болса, онда оның деңгейлері – қыс, көктем, жаз, күз.  
Салыстыралатын топ саны – 4.
- Фактор – емдеу тәсілдерінің түрі болса, онда оның деңгейлері: стандартты әдіспен емдеу, жаңа емдеу түрі және плацебо (бақылау тобы) болуы мүмкін.  
Салыстырылатын топ саны – 3.
- Дисперсиялық талдау жүргізу үшін сапалық белгілер (жыныс, профессия) де, сандық белгілер де (инъекция саны, аурулар саны) қолданылады.



# Дисперсиялық талдаудың негізгі идеясы:

Таңдама дисперсияны екі  
компонентке бөлу:

Факторлық дисперсия

Жалпы ортаға қарасты,  
топтардың орта мәндерінің  
шашырауын сипаттайтын  
факторлық  
дисперсияны *топаралық  
дисперсия* деп атайды.

Қалдық дисперсия.

Топтардың түзетілген  
таңдама дисперсиялары  
үшін орта арифметикалық  
мән болып табылатын  
қалдық  
дисперсияны *топшілік  
дисперсия* деп атайды.

Топтық орта мәндердің теңдігі жөніндегі нөлдік жорамалда топаралық дисперсия топшілік дисперсияға ұқсас. Егер салыстырылып отырған топтар арасында айырмашылық бар болса, онда топаралық дисперсия топшілік дисперсиядан үлкен болады. Фишер критерийі осы екі дисперсияның қатынасына негізделген.

ANOVA үшін Фишер критерийінің F статистикасы топаралық дисперсияның топшілік дисперсияға қатынасы арқылы анықталады.

$$F = \frac{MS_{\text{факт}}}{MS_{\text{калд}}}$$

F статистикасы (k- 1) және (n-k) еркіндік дәрежелеріне сәйкес келетін Фишер таралуына бағынады

**Факторлық  
дисперсия:**

$$MS_{\text{факт}} = r \cdot S_{\bar{x}}^2$$

Мұндағы  $S_{\bar{x}}^2$  – k орта мәндерден тұратын таңдаманың  
( $\bar{x}_{\text{топ1}}, \bar{x}_{\text{топ2}}, \dots, \bar{x}_{\text{топk}}$ )  
таңдама дисперсиясы.

$$S_{\bar{x}}^2 = \frac{1}{k-1} \sum_{i=1}^r (\bar{x}_{\text{mon}j} - \bar{x})^2$$

$$\bar{x} = \frac{1}{k} (\bar{x}_{\text{mon}1} + \bar{x}_{\text{mon}2} + \dots + \bar{x}_{\text{mon}k})$$

## Қалдық дисперсияны есептеу формуласы

$$MS_{\text{қалд}} = \frac{1}{k} (S_1^2 + S_2^2 + \dots + S_k^2)$$

мұндағы  $S_i^2$  –  $i$ -ші таңдаманың дисперсиясы

$$S_i^2 = \frac{1}{r-1} \sum_{i=1}^r (x_{ik} - \bar{x}_{\text{мон } j})^2$$

$MS_{\text{факт}}$  үшін еркіндік дәрежесінің саны  $(k-1)$ , мұндағы  $k$ -топтар саны.

$MS_{\text{алд}}$  үшін еркіндік дәрежесінің саны  $k \cdot (r-1)$ , мұндағы  $r$ -әр топтағы мәндер саны,  $k$  - топтар саны,

## Бірфакторлық дисперсиялық талдау

*Дисперсиялық талдауды жүргізу әдісі:*

1. Нөлдік және балама жорамалдарды құрамыз:

$H_0$ : топтық бас орта мәндер тең, және таңдама орталар арасындағы айырмашылықтар кездейсоқ, фактор оларға ықпал етпейді.

$H_1$ : таңдама орталар арасындағы айырмашылықтар кездейсоқ емес және оларға фактор ықпал етеді.

2.  $\alpha$  мәнділік деңгейі беріледі ( мысалы,  $\alpha=0,05$  немесе  $\alpha=0,01$ ).

### 3. Есептеледі $MS_{\text{факт}}$ және $MS_{\text{калд}}$

Егер,  $MS_{\text{факт}} \leq MS_{\text{калд}}$  онда нөлдік жорамал қабылданады.

Егер,  $MS_{\text{факт}} > MS_{\text{калд}}$  онда  $F = \frac{MS_{\text{факт}}}{MS_{\text{калд}}}$   
Фишер статистикасы есептеледі.

4.  $F_{\text{тәж}}$  есептегеннен кейін,  $F_{\text{сыни}}$  кесте бойынша Фишер таралуының сыни мәндерін табады. Ол  $k-1$  және  $k(r-1)$  еркіндік дәрежелерінің сандарына сәйкес келуі керек.

5.  $F_{\text{тәж}}$  және  $F_{\text{сыни}}$  салыстырылады.

Егер  $F_{\text{тәж}} < F_{\text{сыни}}$ , онда берілген мәнділік деңгейінде  $H_0$  нөлдік жорамалы қабылданады және фактор орта мәнге ықпал етпейді деген қорытынды жасалынады.

Егер  $F_{\text{тәж}} > F_{\text{сыни}}$ , онда нөлдік жорамал жоққа шығарылады және фактор ықпалы маңызды деп танылады.

Параметрлік емес критерийлер бас жиынтықтың таралу түріне тәуелсіз, берілген жиынтықтың варианттары мен олардың жиіліктеріне ғана тәуелді функциялар болып табылады.

Параметрлік емес критерийлер параметрлік критерийлер үшін қажетті болып табылатын таралудың кейбір параметрлерін есептеуді талап етпейді.

Сондықтан параметрлік емес критерийлерді және параметрлік емес статистика әдістерін *параметрден бос* немесе *еркін таралған* деп атайды.

## Параметрлік емес критерийлерді қолданудың тиімділігі мен мүмкіндіктері:

зерттелетін жиынтықтың таралу түрі белгісіз, бұл көбіне көлемі аз жиынтықтармен жұмыс істегенде мәнді;

- сандық және сапалық белгілермен жұмыс істеуге мүмкіндік береді;

- таңдама орта және таңдама ортаның стандартты қатесін есептеу талап етілмейді;

-зерттеліп отырған жиынтықтар арасында айырмашылықтардың бар немесе жоқ екендігін анықтауға, егер бар болса олардың кездейсоқ немесе заңдылық екендігін тағайындауға мүмкіндік береді;

-зерттелетін құбылыстар немесе белгілер арасындағы байланысты немесе тәуелділікті анықтауға мүмкіндік береді.



Крускал Уоллис критерийі -дисперсиялық талдаудың бір факторлы параметрлік емес ұқсастығы.

Ол салыстыру үшін үш немесе оданда көп таңдамаларды қолданады және нөлдік гипотезаны тексереді,соған сәйкес медиананың ортасынан алынған таңдамалар қабылданады.

Крускал Уоллис критерийі бойынша интерпретация параметрлік бір факторлы дисперсиялық анализге сәйкес келеді,бұл критерий орташа көрсеткіштерге қарағанда рангтерге негізделген.

1. Нөлдік және альтернативті гипотезаны анықтау.

$H(0)$ - әр топ популяцияда бірдей таралған.

$H(1)$ - әр топ популяцияда бірдей таралмаған.

2. Бір-біріне ұқсас екі таңдамалардың бірін таңдау.

3.  $H(0)$ – болжамға жауап беретін статистикалық критерийді анықтау.Әрбір топтағы рангтерді есептеп  $n$  белгілерін ранжерлеу: бұл белгілер  $R, \dots, R^k$

Статистикалық критерий мына формула бойынша:

$$H = \frac{12}{n(n+1)} \sum \frac{R_i^2}{n_i} - 3(n+1)$$

4. F - критерийін болжамдық белгілермен салыстыру.

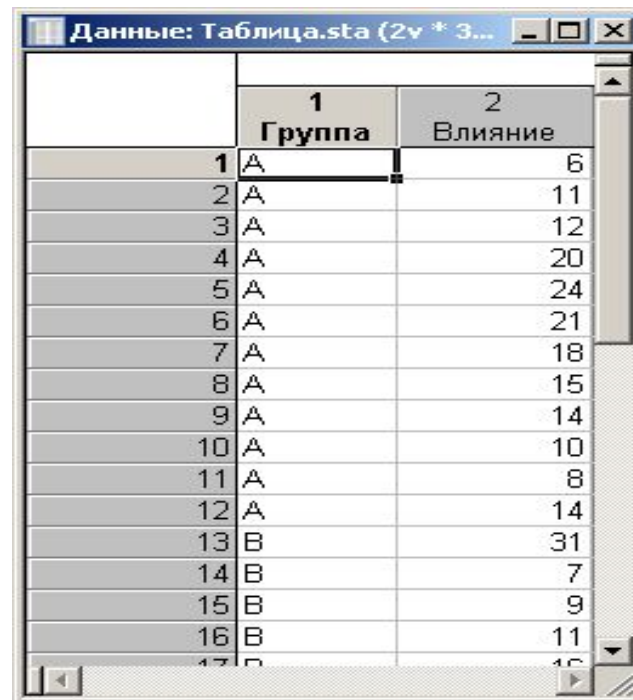
5. Нәтиже мен F жиілікті интерпретациялау.

- Егер нәтиже статистикалық дәлелденген болса жиілікті интерпретациялаймыз. Екі таңдамалы параметрлік емес критерий қолданыңыз, көп ретті тестті талдау үшін коррекциялаңыз. Әр бір топтағы СИ-ын медиана үшін есептейміз. Бір факторлы ANOVA әрбір топ бір факторға сәйкес келсе және тәуелсіз болғанда ғана қолданылады. Зерттеу жұмысының жоспары күрделі болса, онда ANOVA-ның басқа түрін қолдануға болады.

Мысалы: Зерттеу жұмысы барысында X пациенттерге препараттың әсері тексерілді.

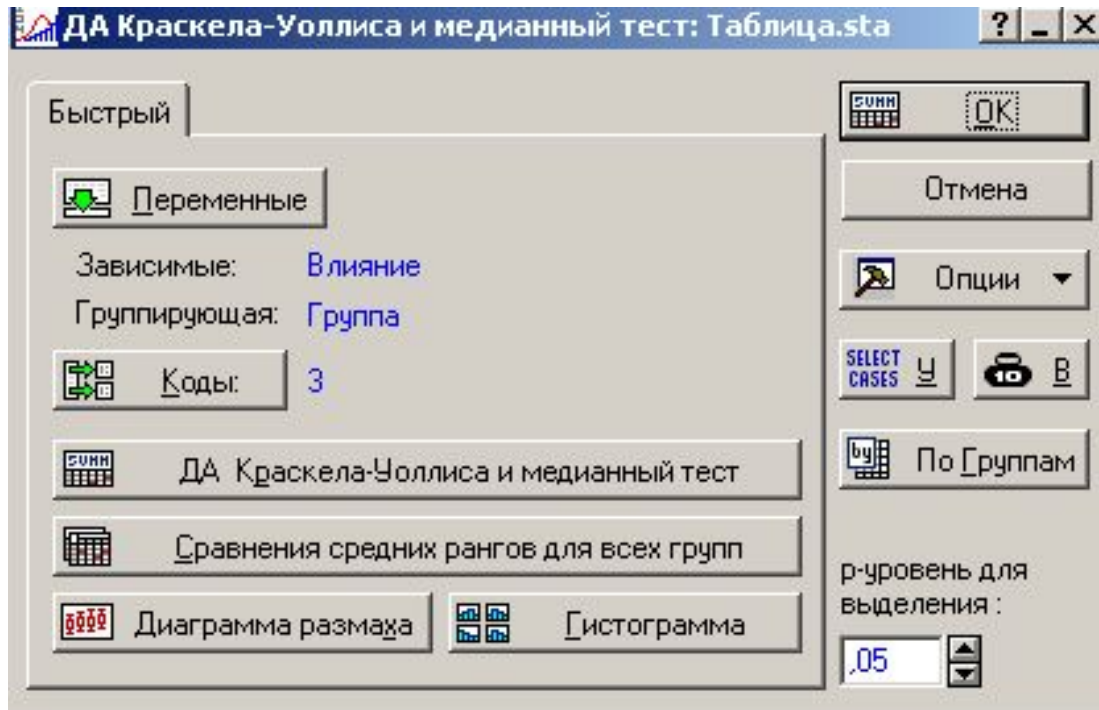
Белгілерге байланысты у көлемі бірдей 3 топқа бөленеді (a,b,c).

Зерттеу жұмысының нәтижесі кесте арқылы берілген.



	1 Группа	2 Влияние
1	A	6
2	A	11
3	A	12
4	A	20
5	A	24
6	A	21
7	A	18
8	A	15
9	A	14
10	A	10
11	A	8
12	A	14
13	B	31
14	B	7
15	B	9
16	B	11
17	B	15

Параметрлік емес статистика командасын анализ мәзірінен параметрлік емес статистика модулін суреттеу үшін таңдаймыз. Әрі қарай бәрнеше тәуелсіз топтарды салыстыру үшін таңдаймыз және Ок батырмасын басып, Крускал Уоллис терезесін ашамыз. Выбор переменных диалогтық терезесін ашу үшін выбор переменных ашамыз. Группа → коды ауыспалы топтарға код таңдаймыз. Диалогты терезесі Крускал Уоллис :



Диалогтық терезесі экраннан шығады.Бұл диалогтық терезеде ОК басып,анализ жасаймыз.

Ранговый ДА Краскела-Уоллиса; Влияние (Таблица. sta)				
Груп. (независ.) переменная: Группа				
Кр.Краскела-Уоллиса: $H(2, N=36) = 13,84438$ $p = ,0010$				
Зависим.: Влияние	Код	Допуст N	Сумма Рангов	Среднее Ранг
<b>A</b>	4	12	139,0000	11,58333
<b>B</b>	6	12	200,0000	16,66667
<b>C</b>	7	12	327,0000	27,25000

Біз К.Уоллис критерийінің жоғары мәнді екенін көреміз.( $P=0,01$ ).Әртүрлі экспериментті топтардың бір-бірімен айырмашылықты табамыз.К.Уоллис бағытталған дисперсиялық анализ деп атайды.

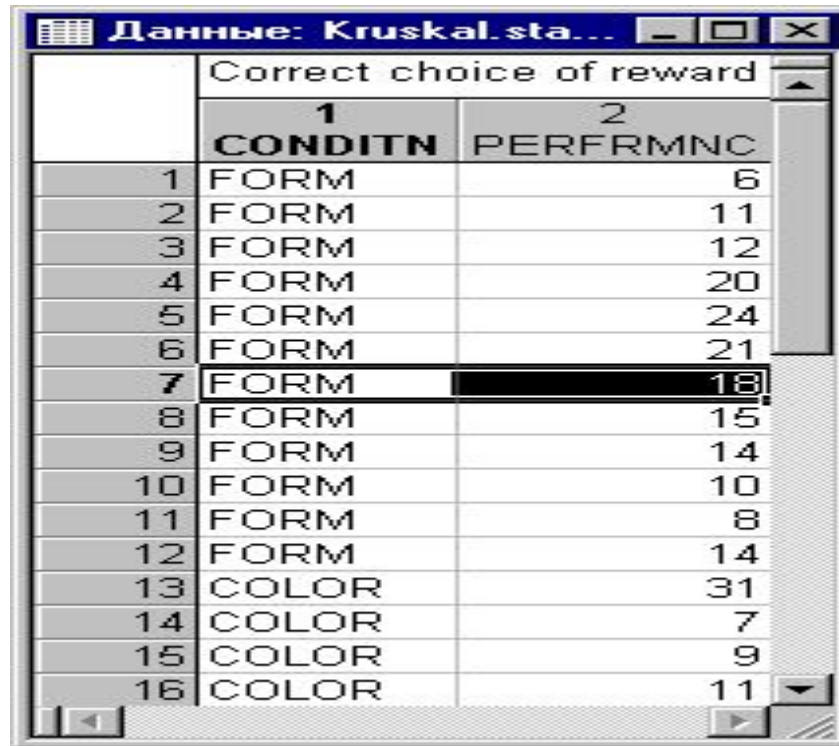
Рангтердің бағасы (әрбір топтарға )нәтижелер кестесіндегі оң жақ бағасында көрсетілген.

Үлкен рангтық баға ( жақсы әсерлі препарат) С тобына қатысты.

Төменгі бағалы ранг (нашар әсер беретін препарат) А тобына қатысты.

Мысал: К. Уоллис дисперциялық анализі және медианалық тест.

Бұл тесттер алтернативті бір факторлы, топшілік Anova.



	Correct choice of reward	
	1	2
	CONDITN	PERFRMNC
1	FORM	6
2	FORM	11
3	FORM	12
4	FORM	20
5	FORM	24
6	FORM	21
7	FORM	18
8	FORM	15
9	FORM	14
10	FORM	10
11	FORM	8
12	FORM	14
13	COLOR	31
14	COLOR	7
15	COLOR	9
16	COLOR	11

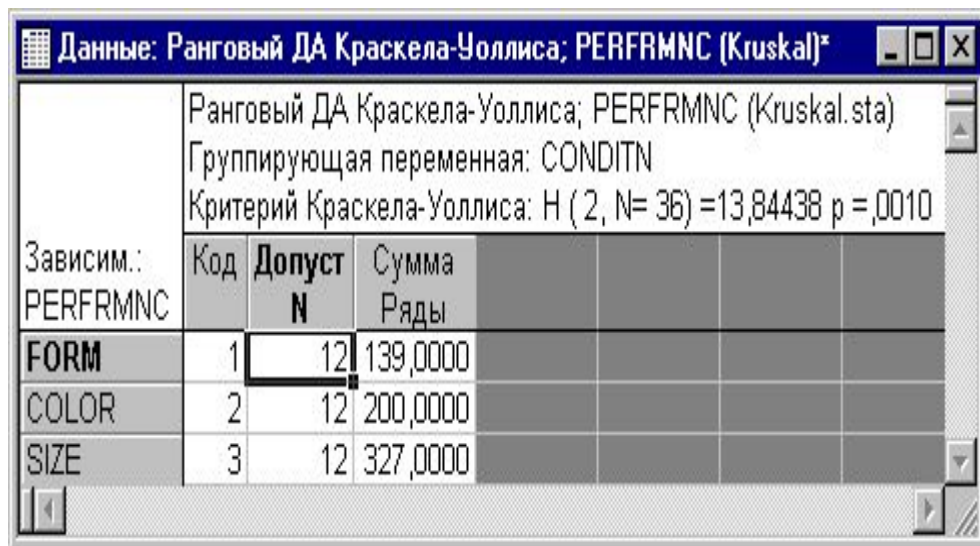
Мысал: (жасанды мәліметтерге және Naus бағдарламасына бағытталған. Бұл мәліметтер кейдейсоқ экспериментті 3 топтарға жазылған. Кішкентай балаларға жасалған зерттеу жұмыстарынан алынған әрбір балаларға жұп тесттер арқылы ұсынылған. Балалардың міндеті дұрыс тыңдап, сыйлық алу болды.

1 топтарға арналған тесттік формасы (1 топ форма – 1 Form)

2 түс (2 топ- түс – 2 Color)

3 өлшем (3- өлшем – Size) заттар ұсынылады. Тәуелді болжам әрбір балаға сыйлық алу үшін міндеттемелер жиналды.

Рангтық ДА К. Уоллис нәтижесі бірінші кесте, медианалық тесттің нәтижесі екінші кестеде көрсетіледі.



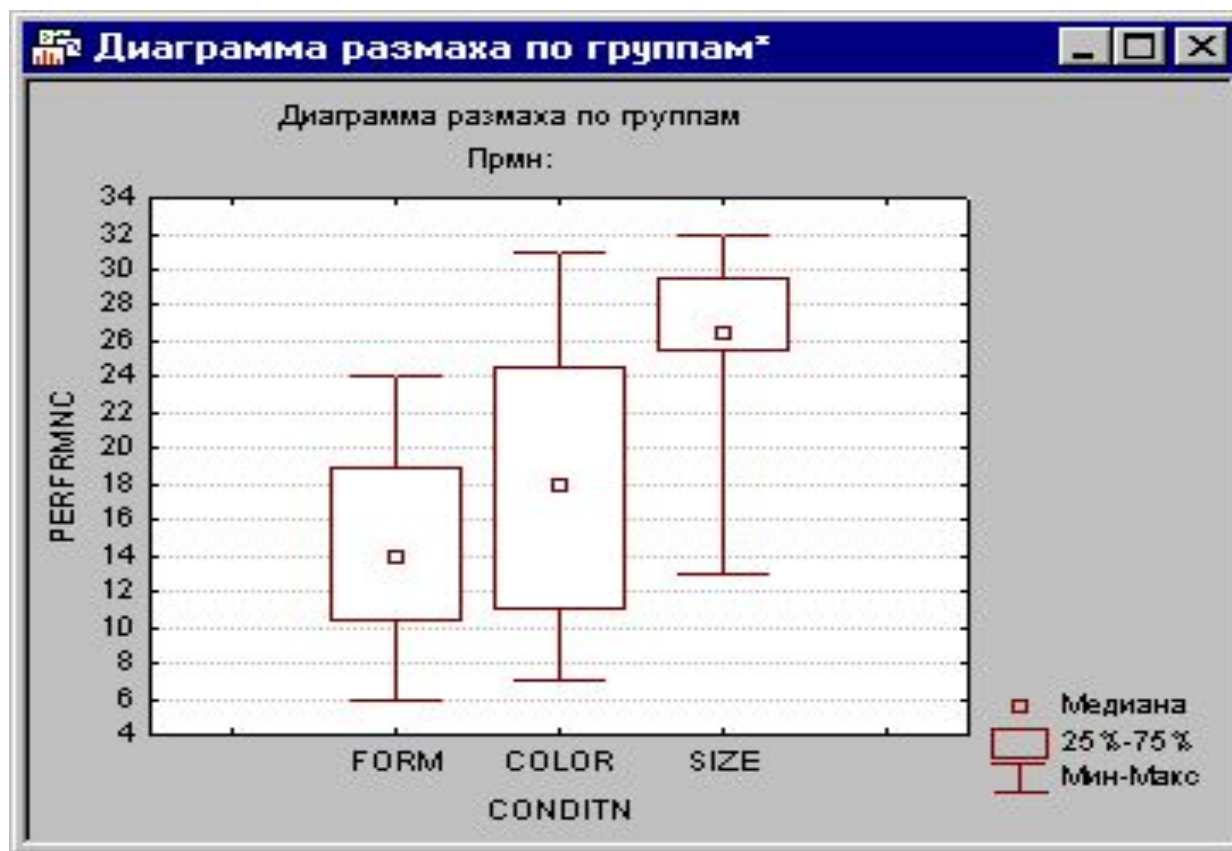
Данные: Ранговый ДА Краскела-Уоллиса; PERFRMNC (Kruskal)\*

Ранговый ДА Краскела-Уоллиса; PERFRMNC (Kruskal.sta)  
Группирующая переменная: CONDITN  
Критерий Краскела-Уоллиса:  $H(2, N=36) = 13,84438$   $p = ,0010$

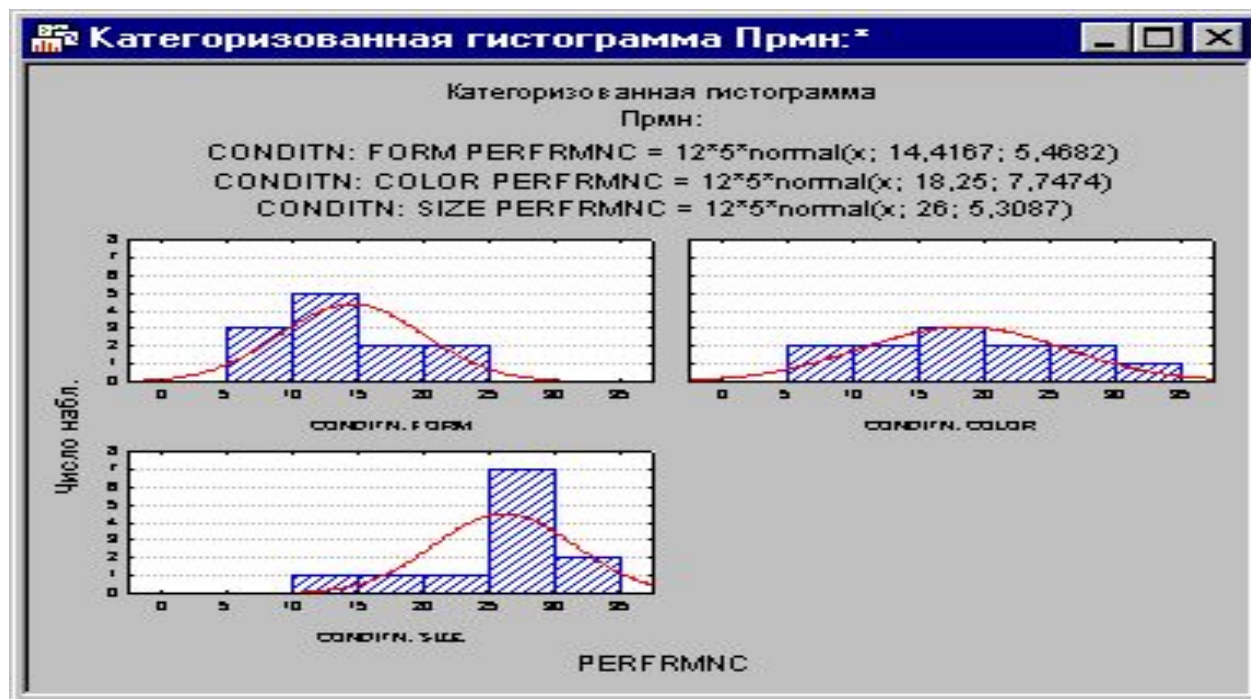
Зависим.:	Код	Допуст N	Сумма Ряды				
PERFRMNC							
FORM	1	12	139,0000				
COLOR	2	12	200,0000				
SIZE	3	12	327,0000				

Медианалық тесттің нәтижесі медианалық критерий нәтижесі мәнді, бірақ маңыздылығы аз.

## Нәтижелердің графикалық көрсетілуі.



Форма тесттік орындау басқаларға қарағанда тиімді екені анық, медиана саны бұл жағдайда басқаларға



Бұл кесте басқа жағдайларға қарағанда Форма тобында «жақсы» орындалғанын дәлелдей түседі. Ең нашар орындалғаны Өлшем тобына арналған кестеде көруге болады. Мұнда балалар Форма тобын оңай ажырата алады деп тұжырымдауға болады.



## Қорытынды.

Қорыта келгенде дисперсиялық талдау- бұл екіден артық топтардың орта мәндерін салыстыру үшін, яғни бірнеше тәуелсіз топтардың бір бас жиынтыққа жататындығын немесе жатпайтындығын анықтау үшін қолданылатын талдау әдісі. Дисперсиялық талдауды алғаш рет ағылшын ғалымы, математик және генетик Рональд Фишер 1938 ж. дисперсиялық талдауға мынадай анықтама берген: Дисперсиялық талдау-«бір себептерден болатын дисперсияны екінші бір себептерден болатын дисперсиядан айыру». Крускал Уоллис критерийі - дисперсиялық талдаудың бір факторлы параметрлік емес ұқсастығын тексеретін әдіс.

## Пайдаланылған әдебиеттер.

Қазақ тіліндегі негізгі:

1. Ахметқазиев А.А., Кельтенова Р.Т. Математикалық статистика, Алматы «Экономика», 2002.

қосымша:

2. Бектаев Қ. Ықтималдықтар теориясы және математикалық статистика. Алматы: «Рауан», 1991.

3. Интернет желісі:

«[http://kazlive.com/referats/kazaksha\\_ref/stan\\_sert\\_metr/bir\\_faktorly\\_dispersijaly\\_taldau.doc](http://kazlive.com/referats/kazaksha_ref/stan_sert_metr/bir_faktorly_dispersijaly_taldau.doc)»,

«<http://www.studfiles.ru>»

«[statistica.ru](http://statistica.ru)»