

با سلام و احترام،

## فصل بیستم و ششم: آزمون ناپارامتریک

قبل از توضیح این آزمون بهتر است تعریفی گذرا از داده‌های پارامتری و غیرپارامتری داشته باشیم:

داده‌های پارامتری، داده‌های اندازه‌گیری شده هستند و بنا به فرض آزمون‌های آماری پارامتری، توزیع این داده‌ها نرمال یا نزدیک به نرمال است. آزمون‌های پارامتری در مورد داده‌های مقیاس فاصله‌ای و مقیاس نسبی به کار می‌روند مانند آزمون  $t$  و  $F$ ، بنابراین داده‌های اسمی (شمارش فراوانی‌ها) و ترتیبی (دسته‌بندی) برای استفاده از آزمون‌های پارامتری مناسب نیستند. داده‌های ناپارامتری، داده‌هایی قابل شمارش‌اند یا رتبه‌بندی می‌شوند. آزمون‌های ناپارامتری که گاهی به عنوان آزمون‌های مستقل از توزیع شناخته می‌شوند به پیش فرض نرمال بودن توزیع جامعه‌ها استوار نیستند یعنی فرض نرمال بودن جامعه را لازم ندارند.

چند آزمون ناپارامتریک این نرم افزار در دسترس است که عبارتند از:

✱ آزمون مربع کای (خی دو یا کی دو) تک متغیره

✱ آزمون دو جمله‌ای

✱ آزمون **Runs**

✱ آزمون تک نمونه‌ای کولموگروف - اسمیرنوف

✱ آزمون با دو نمونه مستقل

✱ آزمون با چند نمونه مستقل

✱ آزمون با دو نمونه وابسته

✱ آزمون با نمونه‌های وابسته

آزمون مربع کای (خی دو یا کی دو) تک متغیره

دستورالعمل **chi - Square Test** یک متغیره را در مقوله‌ها جدول‌بندی می‌کند و فرضیه‌های مبنی بر اینکه فراوانی‌های مشاهده شده از فراوانی‌های مورد انتظارشان تفاوت نمی‌کنند را مورد آزمون قرار می‌دهد.

آزمون مربع کای:

- \* تمام مقوله‌های متغیر آزمون را شامل می‌شود، یا آزمون را به دامنه‌ای معین محدود می‌نماید.
- \* مقادیر استاندارد یا سفارشی شده مورد انتظار را به کار می‌برد.
- \* آماره‌های توصیفی و یا چارک‌ها را در متغیر آزمون به دست می‌آورد.

## آزمون عدم وابستگی

در یک برنامه زمان‌بندی بیمارستانی بزرگ برای ترخیص بیماران فرض بر این شده است که بیماران با نرخ ثابتی در هفته بیمارستان را ترک می‌کنند. مسئولین بیمارستان می‌خواهند بدانند که آیا تعداد ترخیصی‌ها با روزهای هفته متغیر است یا نه. با آزمون **chi – Square** فرضیه رفتن بیماران از بیمارستان با سرعت ثابت را بررسی نمایید.

## آماده کردن داده‌ها

هر مورد یک روز در هفته می‌باشد، و برای اجرای آزمون مربع کای، باید ابتدا موارد را با فراوانی ترخیص بیمار وزن نمایید.

۱. برای وزن‌دهی موارد، مسیر **Data > Weight Cases** را از نوار منوی **Data Editor** برگزینید.
۲. گزینه **Weight Cases by** را فعال کنید.
۳. گزینه **Average Daily Discharges** را به عنوان متغیر فراوانی انتخاب نمایید.
۴. **Ok** را کلیک کنید.

موارد اکنون با فراوانی ترخیص بیمار وزن می‌شوند.

## آغاز تحلیل

۱. مسیر **Analyze > Nonparametric Tests > Legacy Dialogs > Chi – square** را از منوی اصلی برگزینید.
۲. **Day of the Week** را به عنوان متغیر آزمون برگزینید.
۳. **Ok** را کلیک کنید.

در کادر محاوره فوق بخشی با عنوان **Expected values** وجود دارد. که در آن می‌توان فراوانی‌های مورد انتظار را در دو حالت تنظیم کرد:


**All Categories equal**: با انتخاب این گزینه، فراوانی‌های هر طبقه برابر در نظر گرفته می‌شود. در این حالت، فراوانی‌های مورد انتظار برای هر طبقه، از تقسیم تعداد کل فراوانی‌ها (**N**) بر تعداد طبقات متغیر به دست می‌آید.

**Values:** با انتخاب این گزینه، می‌توانیم بر اساس یک مبنای نظری، فراوانی‌های مورد انتظار هر طبقه را به صورت دستی و نابرابر تعریف کنیم. یعنی فراوانی‌های تمام طبقات را که انتظار داریم فراوانی‌های آنها بیشتر و یا کمتر از بقیه باشد، مشخص می‌کنیم.

## جدول فراوانی مربع کای

در اینجا، فراوانی مشاهده شده (**Observed**) برای هر سطر به سادگی میانگین تعداد بیماران ترخیصی در روز در کل سال می‌باشد. مثلاً در سال گذشته، بیمارستان تعداد میانگین **589** بیمار را در هفته، **44** بیمار را در یکشنبه، **78** بیمار را در دوشنبه و ... ترخیص کرد (**a**).

فراوانی مورد انتظار (**Expected**) مربوط به هر سطر برابر مجموع فراوانی‌های مشاهده شده تقسیم بر تعداد سطرها در جدول می‌باشد. در این مثال، **589** ترخیص در هفته می‌باشد، که حدود **84** ترخیص را در روز نتیجه می‌دهد (**b**).



	a	Day of the Week	b	
	Observed N		Expected N	Residual
Sunday	44		84.1	-40.1
Monday	78		84.1	-6.1
Tuesday	90		84.1	5.9
Wednesday	94		84.1	9.9
Thursday	89		84.1	4.9
Friday	110		84.1	25.9
Saturday	84		84.1	-1.1
Total	589			

در آخر، باقی‌مانده (**Residual**) برابر با فراوانی مشاهده شده منهای مقدار مورد انتظار (ستون دوم) می‌باشد. جدول نشان می‌دهد که یکشنبه خیلی کمتر، و جمعه خیلی بیشتر (از فرضیه «هر روز مساوی است») انتظار می‌رود.

## جدول آزمون مربع کای

آماره به دست آمده مربع کای برابر **29.389** می‌باشد. این مقدار با مربع گرفتن از باقی‌مانده هر روز، تقسیم بر مقدار مورد انتظار، و جمع آن در سرتاسر روزها محاسبه شده است.

عبارت **df** درجه آزادی را نشان می‌دهد. در یک آزمون مربع کای، **df** تعداد مقادیر مورد انتظار است که می‌تواند قبل از اینکه باقی‌مانده کاملاً تعیین شود تغییر نماید. در مورد یک آزمون مربع کای تک موردی، **df** برابر تعداد سطرها منهای **1** است.

**Asymp.sig** برآورد احتمالی حاصله از یک مقدار مربع کای بزرگتر یا مساوی **29.389** می باشد. به شرطی که بیماران به صورت یکنواخت در طول هفته ترخیص شوند. سطح معنی داری کوچک به این معنی است که میانگین نرخ ترخیص بیماران در روزهای هفته متفاوت است.

	Day of the Week
Chi-Square <sup>a</sup>	29.389
df	6
Asymp. Sig.	.000

## بررسی یک دامنه معین

در حالت پیش فرض، دستورالعمل آزمون مربع کای بر اساس مقادیر مهمتر متغیر آزمون، فراوانی ها را ایجاد کرده و یک مقدار مورد انتظار را محاسبه می نماید. با این وجود، ممکن است بخواهید که دامنه آزمون را به یک مجموعه مقادیر معین محدود نمایید. همانطور که مثال بعد نشان می دهد، دستورالعمل به آسانی برای این به کار می رود. بیمارستان در ادامه تحلیل می خواهد بداند که آیا می توان در یک برنامه بندی زمانی بیماران را از دوشنبه تا جمعه با نرخ ثابت روزانه ترخیص نمود.

## آغاز تحلیل

۱. برای بازگشت به تحلیل، کادر محاوره **chi – Square Test** را فراخوانی نمایید.
۲. گزینه **Use Specified range** را انتخاب کنید.
۳. **2** را برای کمترین مقدار و **6** را برای بیشترین مقدار تایپ کنید.
۴. **Ok** را کلیک کنید.

## جدول فراوانی روزهای هفته

دامنه آزمون به دوشنبه تا جمعه محدود شده است. به طور متوسط، حدود **92** بیمار در روزهای کاری از بیمارستان ترخیص می شوند **(a)**. ستون **Residual** نشان می دهد که دوشنبه ها سبک تر و جمعه ها سنگین تر هستند. البته، فقط با نگاه کردن به اینها نمی توانیم بگوییم که این اختلافات معنی دار هستند.



	Day of the Week <sup>a</sup>			
	Category	Observed N	Expected N	Residual
1	Monday	78	92.2	-14.2
2	Tuesday	90	92.2	-2.2
3	Wednesday	94	92.2	1.8
4	Thursday	89	92.2	-3.2
5	Friday	110	92.2	17.8
Total		461		

### آزمون مربع کای روزهای هفته

با درجه آزادی ۴، خی دو به دست آمده (۵/۸۲۲) به هیچ سطح معنی داری قابل قبولی نزدیک نیست در نتیجه ما نمی توانیم فرضیه ای را که می گوید بیماران با نرخ حدود ۹۲ نفر در هفته ترخیص می شوند رد نماییم.

Day of the Week	
Chi-Square <sup>a</sup>	5.822
df	4
Asymp. Sig.	.213

### سفارشی کردن مقادیر مورد انتظار

آزمون مربع کای معمولاً برای بررسی فراوانی های مشاهده شده در مقابل یک مقدار مورد انتظار واحد که برابر همه سطرها می باشد به کار می رود. با این وجود، توزیع مقادیر ممکن است از آن الگو تبعیت نکند.

مثلاً در علم ژنتیک، انتظار دارید که خصوصیات غالب را در ۷۵% مردم و مغلوب را در ۲۵% باقی مانده ببینید. دستورالعمل آزمون مربع کای این امکان را فراهم می آورد که یک مجموعه سفارشی از مقادیر مورد انتظار را تعیین نمایید، از این راه یک دامنه وسیع از مدل ها بررسی می شوند.

یک تولید کننده پوشاک سعی دارد از پست درجه ۱ استفاده نماید، تا سریع تر جواب بگیرد. مشتریان سفارش دهنده تعداد هفته هایی را که ارسال سفارشی طول می کشد ثبت می نمایند. با استفاده از آزمون مربع کای تعیین کنید که آیا درصد سفارشات در هفته با دو روش متفاوت هستند

## آماده کردن داده‌ها

داده‌ها به صورت هفتگی جدول‌بندی شده‌اند و باید ابتدا با فراوانی پاسخ پست درجه اول وزن‌دهی شوند.

۱. برای وزن‌دهی داده‌ها، مسیر **Data > Weight Cases** را از نوار منوی اصلی در **Data Editor** برگزینید.
۲. گزینه **Weight Cases by** را انتخاب کنید.
۳. **First Class Mail** را به عنوان متغیر فراوانی برگزینید.
۴. **Ok** را کلیک کنید.

اکنون داده‌ها وزن شده‌اند و آماده تحلیل هستند.

## شروع تحلیل

۱. برای آغاز تحلیل، مسیر **Analyze > Nonparametric Tests > Legacy Dialogs > Chi – square** را از منوی اصلی برگزینید.
۲. **Week of Response** را به عنوان متغیر آزمون برگزینید.
۳. **Values** را در مجموعه **Expected Values** انتخاب نمایید.
۴. عدد 6 را به عنوان اولین مقدار مورد انتظار تایپ کرده و **Add** را کلیک کنید.
۵. این رویه را تکرار کنید و مقادیر **15.1, 18, 12, 11.5, 9.8, 7, 6.1, 5.5, 3.9, 2.1** و 2 را اضافه نمایید.
۶. بعد از وارد کردن ۱۲، دکمه **Ok** را کلیک کنید.

## تعیین انطباق بین مشاهدات و پاسخ‌های مورد انتظار

- درصد‌های پاسخ مشاهده شده در ستون **Observed N** ظاهر می‌گردد. به دلیل استفاده این متغیر به عنوان متغیر وزن‌دهی، اینها مقادیر واقعی متغیر **First Class Mail** می‌باشند.
- درصد‌های پاسخ مورد انتظار در ستون **Expected N** ظاهر می‌شود. توجه نمایید که مقادیر در این ستون صرفاً درصد‌های تاریخی مربوط به پست معمولی هستند که اینها را در کادر محاوره **chi – Square Test** وارد کردید.
- تفاوت بین درصد‌های پاسخ پست درجه ۱ و معمولی در ستون **Residual** ظاهر می‌شود.

	Observed N	Expected N	Residual
1	10	6.0	4.0
2	22	15.1	6.9
3	14	18.0	-4.0
4	10	12.0	-2.0
5	9	11.5	-2.5
6	8	9.8	-1.8
7	7	7.0	.0
8	10	6.1	3.9
9	4	5.5	-1.5
10	2	3.9	-1.9
11	2	2.1	-.1
12	1	2.0	-1.0
Total	99		

شرکت امیدوار است که پست درجه ۱ سریع تر پاسخ مشتریان را بدهد. در دو هفته اول، این اختلاف معادل چهار و هفت دهم درصد می باشد. سؤال این است که آیا اختلاف بین دو توزیع کلی به اندازه کافی بزرگ می باشد تا یک اختلاف ایجاد نماید.

در ۱۱ درجه آزادی، آماره به دست آمده مربع کای (12.249) از نظر آماری معنی دار نیست. این ارتقاء خاص منجر به زمان های پاسخی نمی شود که به طور معنی داری / از پست معمولی استاندارد تفاوت دارد.



ادامه فصل را در مرجع کاربردی SPSS 20 (45) دنبال نمایید.



**Telegram.me/iepnu**  
 کانال تخصصی مهندسی صنایع دانشگاه پیام نور