

## فصل بیستم و یکم: تک متغیره GLM

با دستورالعمل **GLM Univariate** می‌توانید مقدار یک متغیر کمی وابسته را بر اساس رابطه‌اش با پیشگوهای مقوله‌ای و کمی مدل کرد.

این آزمون که تعمیم یافته آزمون تحلیل واریانس یک طرفه می‌باشد، از جمله آزمون‌های پارامتری است که کاربرد زیادی در تحلیل آماری دارد. آزمون تحلیل واریانس تک متغیره در کنار روش‌های تحلیل واریانس چند متغیره به **روش‌های تحلیل واریانس فاکتوریل** معروفند. دلیل اطلاق نام فاکتوریل به این آزمون‌ها، آن است که در آن، از چند متغیر مستقل استفاده می‌کنیم و قصد داریم ببینیم کدام متغیر مستقل بیشترین اثر را بر متغیر وابسته دارد. طبیعی است که چنین کاری را نمی‌توانیم با آزمایش‌های جداگانه انجام دهیم و بر اساس نتایج هر آزمایش پی ببریم که متغیرهای مستقل در تعامل با هم چگونه عمل می‌کنند و در ترکیب خاصی اثر منحصر به فردی ایجاد می‌کند.

بنابراین، در این آزمون، برخلاف آزمون تحلیل واریانس یک طرفه که با یک متغیر مستقل طبقه‌بندی شده (اسمی / ترتیبی) و یک متغیر وابسته (فاصله‌ای / نسبی) مواجه بودیم، با حالتی روبرو هستیم که متغیر مستقل ما شامل دو یا چند متغیر است، اما باز با یک متغیر وابسته سروکار داریم. از این آزمون به خاطر استفاده از دو یا چند متغیر مستقل، آزمون تحلیل واریانس دو طرفه نیز نام برده می‌شود. به عبارتی، این روش بر پایه روش مدل عمومی خطی است که در آن فرض بر این می‌باشد که عامل‌ها و کوواریت‌ها یک رابطه خطی با متغیر وابسته دارد.

به طور خلاصه، در این آزمون هدف این است که نه تنها تأثیر جداگانه دو یا چند متغیر مستقل کیفی بر یک متغیر وابسته کمی آزمون شود بلکه تأثیر همزمان این دو متغیر بر متغیر وابسته نیز مورد نظر است.

مدل تک متغیره GLM

دستورالعمل تک متغیره GLM بر اساس دستورالعمل General Linear Model (مدل خطی عمومی) بنا شده است، فرض بر این است که عوامل و متغیرهای تصادفی کمکی (کواریت‌ها) یک رابطه خطی با متغیر وابسته دارند.

به منظور آزمایش فرضیه‌های مربوط به برآوردهای پارامتر، تحلیل تک متغیره، GLM موارد زیر را فرض می‌کند:

- ❖ مقادیر خطا مستقل از هر مقدار دیگر و متغیرهای در مدل می‌باشد. مطالعه کافی طرح عموماً مانع از نقض این فرض می‌شود.
- ❖ پراکندگی خطاها در سرتاسر سلول‌ها ثابت است. وقتی اندازه سلول‌های غیر مساوی وجود داشته باشد این نکته می‌تواند خصوصاً مهم باشد؛ (به عبارت دیگر، تعداد اختلاف مشاهده شده در سرتاسر ترکیبات سطح عامل).
- ❖ خطاها یک توزیع نرمال با یک مقدار میانگین دارند.

## استفاده از تک متغیره GLM برای اجرای یک تحلیل دو عامل واریانس

یک فروشگاه زنجیره‌ای خواروبار فروشی از یک سری مشتریان خود نظرسنجی کرده است. با نتایج به دست آمده و مقدار هزینه‌ای که مشتری در ماه گذشته خرج کرده است، فروشگاه می‌خواهد ببیند که اگر فراوانی مشتریان فروشگاه با مقدار هزینه‌ای که صرف کرده‌اند در ارتباط است، از آن برای کنترل جنسیت مشتری استفاده نماید. از دستورالعمل تک متغیره GLM برای اجرای یک ANOVA دو طرفه (دوراهه) روی مقدار هزینه‌ها استفاده نمایید.

## آغاز تحلیل

۱. مسیر **Analyze > General Linear Model > Univariate** را بروی شروع تحلیل GLM انتخاب نمایید.

۲. Amount Spent را به عنوان متغیر وابسته برگزینید.
۳. Gender و Shopping Style را برای عوامل ثابت (Fixed Factor) برگزینید.
۴. دکمه Plots را کلیک کنید.
۵. Style را به عنوان متغیر محور افقی (Horizontal Axis) انتخاب نمایید. دقت کنید که عناوین متغیر در کادر محاوره‌های فرعی GLM Univariable ظاهر نمی‌شوند.
۶. gender را به عنوان متغیر خطوط جدا کننده (Separate Lines) برگزینید.
۷. روی Add کلیک کنید.
۸. دکمه Continue را کلیک کنید.
۹. Post Hoc را در کادر محاوره GLM Univariate کلیک کنید.
۱۰. Style را به عنوان متغیری که برای آزمون‌های پست هاک باید ایجاد شود در قسمت Post Hoc Tests for برگزینید.
۱۱. گزینه Tukey را در مجموعه Equal Variances Not Assumed فعال نمایید.
۱۲. دکمه Continue را کلیک کنید.
۱۳. دکمه Options را در کادر محاوره GLM Univariate کلیک کنید.
۱۴. gender\* style را به عنوان عبارتی که برای میانگین‌ها باید ظاهر شود انتخاب نمایید.
۱۵. گزینه‌های Descriptive Statistics، Homogeneity tests، Estimates of effect Size و Spread Vs. Level plot را در مجموعه Display فعال نمایید.
۱۶. دکمه Continue را کلیک کنید.
۱۷. Ok را در کادر محاوره GLM Univariate کلیک کنید.

این مراحل یک تحلیل دوراها از واریانس ایجاد می‌نماید. آزمون‌های پست هاک نحوه تفاوت سه سبک خرید را تعیین می‌کنند.

آماره‌های توصیفی

این جدول آماره‌های توصیفی برای ترکیب‌های عوامل در مدل را نشان می‌دهد. سبک خرید بدین صورت است که در حالت میانگین، مشتری‌ها به صورت دو هفتگی مبلغ **\$378.52** را هزینه می‌کنند، در حالی که به صورت هفتگی **\$404.55** و اغلب مبلغ **\$406.76** پرداخت می‌نمایند **(a)**.

از نظر جنسیتی، به طور متوسط، مردان در نمونه، مبلغ **\$430.30** در مقایسه با زنان (**\$365.66**) هزینه می‌کنند **(b)**. سرانجام، ممکن است یک اثر تعاملی (**interaction**) بین جنسیت و سبک خرید باشد، زیرا تفاوت‌های میانگین در مقدار هزینه شده با سبک خرید بین جنسیت‌های مختلف فرق می‌کند **(c)**.

Dependent Variable: Amount spent

Gender	Shopping style	Mean	Std. Deviation	N
Male	Biweekly, in bulk	413.0657	90.86574	35
	Weekly, similar items	440.9647	98.73560	120
	Often, what's on sale	407.7747	69.33334	30
	Total	430.3043	93.47677	185
Female	Biweekly, in bulk	343.9753	100.47207	35
	Weekly, similar items	361.7206	90.46076	102
	Often, what's on sale	405.7269	80.57058	29
	Total	365.6671	92.64058	166
Total	Biweekly, in bulk	378.5210	101.25833	70
	Weekly, similar items	404.5552	102.40440	222
	Often, what's on sale	406.7681	74.47114	59
	Total	399.7352	98.40821	351

مثلاً، مشتریان مرد دو هفتگی مبلغی بیشتر از مشتریان مرد که اغلب (**often**) خرید می‌کنند هزینه می‌نمایند، اما همین نسبت برای زنان برعکس می‌باشد **(a)**.

ستون **N** در جدول نشان می‌دهد که گروه‌های غیر هم اندازه وجود دارند. اغلب مشتریان ترجیح می‌دهند که به صورت هفتگی خرید کنند **(b)**.

انحراف معیار به صورت نسبتاً یکنواخت ظاهر می‌شود، اگر چه برای اطمینان می‌توانید آزمون‌های لون و نمودارها را بررسی نمایید **(c)**.



Dependent Variable: Amount spent

Gender	Shopping style	Mean	Std. Deviation	N
Male	Biweekly, in bulk	413.0657	90.86574	35
	Weekly, similar items	440.9647	98.23860	120
	Often, what's on sale	407.7747	69.33334	30
	Total	430.3043	93.47877	185
Female	Biweekly, in bulk	343.9763	100.47207	35
	Weekly, similar items	361.7205	90.46076	102
	Often, what's on sale	405.7269	80.57050	29
	Total	385.6871	92.64050	166
Total	Biweekly, in bulk	378.5210	101.25839	70
	Weekly, similar items	404.5552	102.48440	222
	Often, what's on sale	406.7681	74.42114	59
	Total	399.7352	98.40621	351

## بررسی یکنواختی واریانس‌ها

این جدول، فرضیه صفری را که می‌گوید واریانس خطا در سرتاسر گروه‌های تعریف شده با ترکیب سطوح عامل ثابت است بررسی می‌نماید از آنجا که مقدار معنی‌داری آزمون، یعنی **0.30**، بزرگتر از **0.05** است، دلیلی ندارد که فرض مساوی بودن واریانس‌ها نقض شود. بنابراین، اختلافات کوچک در گروه انحراف‌های معیار مشاهده شده در جدول آماره‌های توصیفی به خاطر تغییر تصادفی هستند **(a)**.

Dependent Variable: Amount spent

F	df1	df2	Sig.
1.157	5	345	.330

## آزمون‌های پست هاک

آزمون‌های تأثیرات بین موردی به شما در تعیین معنی‌داری یک عامل کمک می‌کند. هر چند، مشخص نمی‌کند که چطور سطوح یک عامل فرق می‌کنند. آزمون‌های پست هاک اختلافات در میانگین‌های مدل پیشگو را برای هر جفت سطوح عامل نشان می‌دهد.

ستون اول آزمون‌های اختلاف پست هاک را نشان می‌دهد **(a)**. دو ستون بعد، جفت سطوح عامل که آزمون شده‌اند را نشان می‌دهد **(b)**. وقتی که سطح معنی‌داری برای اختلاف در مبلغ هزینه شده (**Amount Spent**) مربوط به یک جفت سطوح عامل کمتر از **0.05** باشد، یک علامت \* همراه با اختلاف چاپ می‌شود **(c)**. در این مورد، هیچ تفاوت معنی‌داری در هزینه‌های دو هفتگی، هفتگی یا اغلب (**Often**) مشتریان وجود ندارد.



Dependent Variable: Amount spent

a	b		Mean Difference (I-J)	c		d	
	(I) Shopping style	(J) Shopping style		Std. Error	Sig.	Lower Bound	Upper Bound
Tukey HSD	Biweekly, in bulk	Weekly, similar items	-26.0342	12.611	.099	-55.7191	3.6507
		Often, what's on sale	-28.2471	16.259	.193	-66.5198	10.0256
		Weekly, similar items	26.0342	12.611	.099	-3.6507	55.7191
	Weekly, similar items	Often, what's on sale	-2.2130	13.475	.985	-33.9320	29.5061
		Biweekly, in bulk	28.2471	16.259	.193	-10.0256	66.5198
		Weekly, similar items	2.2130	13.475	.985	-29.5061	33.9320
Tamhane	Biweekly, in bulk	Weekly, similar items	-26.0342	13.921	.180	-59.7559	7.6875
		Often, what's on sale	-28.2471	15.503	.190	-65.7680	9.2737
		Weekly, similar items	26.0342	13.921	.180	-7.6875	59.7559
	Weekly, similar items	Often, what's on sale	-2.2130	11.882	.997	-30.9758	26.5499
		Biweekly, in bulk	28.2471	15.503	.190	-9.2737	65.7680
		Weekly, similar items	2.2130	11.882	.997	-26.5499	30.9758

Based on observed means.

وقتی اندازه‌های گروه نامساوی باشند **Tamhane's T2** عموماً مناسب‌تر از **Tukey's HSD** می‌باشد، اما نتایج در این مورد تا حد زیادی یکسان می‌باشد. این گزینه، آزمون محافظه کاری است که بر اساس مقدار آماره **t**، به مقایسه جفتی گروه‌ها با همدیگر می‌پردازد.

فاصله‌های اطمینان برای **Tamhane's T2** فقط کمی بازتر از **Tukey's HSD** می‌باشد.

چون نتایج این دو آزمون خیلی متفاوت نیستند، می‌توانید نتایج زیر مجموعه‌های یکنواخت را بررسی نمایید که برای **Tukey's HSD** در دسترس هستند اما برای **Tamhane's T2** در دسترس نمی‌باشند **(d)**.

جدول زیر مجموعه نتایج آزمون‌های پست هاک را برمی‌دارد و آنها را در یک شکل قابل تفسیرتر نشان می‌دهد. در ستون‌های زیر مجموعه، سطوح عاملی که تأثیرات، تفاوت معناداری ندارند ظاهر می‌شوند. در این مثال، اولین زیر مجموعه شامل مشتری‌های دو هفتگی، هفتگی یا اغلب **(Often)** می‌باشد **(a)**. همه مشتری‌ها وجود دارند، بنابراین زیر مجموعه‌های دیگر موجود نیستند.

آزمون‌های پست هاک حکایت از آن دارند که تلاش برای جلب مشتریان برای خرید بیشتر بی‌نتیجه است زیرا آنها هزینه قابل توجه بیشتری نخواهند کرد. با این وجود، نتایج آزمون پست هاک برای سطوح عوامل دیگر به



حساب نمی‌آیند، بنابراین صرفنظر کردن از یک تأثیر تعاملی با جنسیت در جدول آماره‌های توصیفی به چشم می‌خورد.

Shopping style	N	a	Subset
Tukey HSD +1 Biweekly, in bulk	70		378.5210
Weekly, similar items	222		404.5552
Often, what's on sale	59		406.7681
Sig.			.116

## میانگین‌های حاشیه‌ای تخمینی

این جدول میانگین‌های حاشیه‌ای تخمینی و خطاهای استاندارد مقادیر هزینه شده (**Amount Spent**) در عامل تشکیل شده از **Gender** و **Shopping Style** را نشان می‌دهد. و برای بررسی تأثیر تعامل ممکن بین این دو عامل کاربرد دارد (a).

در این مثال، یک مشتری آقا که خرید هفتگی انجام می‌دهد انتظار می‌رود مبلغ **\$440.96** خرید انجام دهد، در حالی که مشتری دیگری که خرید به صورت اغلب (**Often**) انجام می‌دهد انتظار می‌رود خریدی معادل **\$407.77** انجام دهد (b).

یک مشتری خانم که خرید او به صورت هفتگی است از او خریدی **\$361.72** انتظار می‌رود، در حالی که خانمی که خریدش به صورت اغلبی (**Often**) است از او معادل **\$405.72** انتظار می‌رود. بنابراین یک فرق بین مشتری‌های **weekly** و **often** (بسته به جنسیت مشتری) وجود دارد (c).

Dependent Variable: Amount spent		95% Confidence Interval			
Gender	Shopping style	Mean	Std. Error	Lower Bound	Upper Bound
Male	Biweekly, in bulk	413.066	15.551	382.479	443.652
	Weekly, similar items	440.965	8.398	424.446	457.483
	Often, what's on sale	407.775	16.797	374.738	440.812
Female	Biweekly, in bulk	343.976	15.551	313.390	374.563
	Weekly, similar items	361.720	9.109	343.604	379.837
	Often, what's on sale	405.723	17.084	372.125	439.329

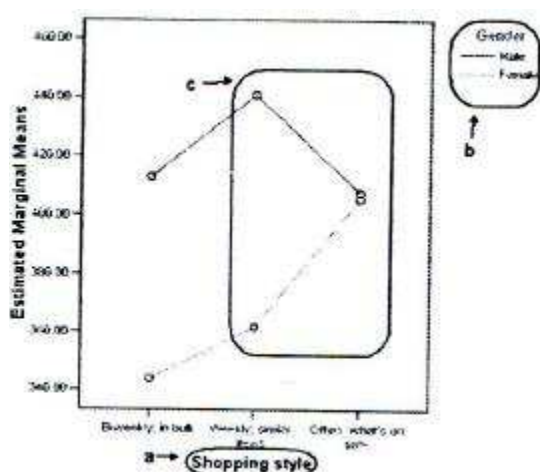
این واقعیت از یک تأثیر تعاملی بین جنسیت و نوع خرید حکایت دارد. اگر هیچ تعاملی وجود نداشت، انتظار داشتید که تفاوت بین انواع خرید برای مشتری‌های آقا و خانم ثابت بماند. تعامل را در نمودارهای تعامل

(Profile) راحت تر می توان مشاهده نمود. نمودر تعامل یک نمایش تصویری از جدول میانگین های حاشیه ای است.

سطوح عامل Shopping Style در محور افقی نشان داده شده اند (a). خطوط جدایش برای هر سطح از جنسیت ایجاد شده اند (b).

به نوبت، سطوح عامل Gender (جنسیت) می تواند در محور افقی، به همراه خطوط جدایش برای هر سطح Shopping Style نمایش داده شود.

در صورتی که هیچ تأثیر تعاملی وجود نداشته باشد، خطوط در جدول موازی می باشند. در عوض، می توانید ببینید که اختلاف در هزینه بین مشتری های weekly و Often برای مشتری های زن بزرگتر می باشد، به طوری که خط مربوط به مشتریان مرد شیب رو به پایین و خط مربوط به مشتریان زن شیب رو به بالا دارد (c).



این یک تأثیر تعاملی قوی است و احتمال اینکه ناشی از اتفاق باشد وجود ندارد، اما شما باید آزمون های تأثیرات بین موردی (between-subject) را برای اثبات معنی دار بودن آن بررسی نمایید.

## آزمون های تأثیرات بین موردی

منظور تحلیل جدول واریانس است. هر ترم در مدل، به علاوه مدل به صورت یکپارچه، از نظر توانایی آن برای تغییر در متغیر وابسته بررسی می شود. توجه کنید که عناوین متغیر در این جدول نمایش داده نشده اند (a). سطح



معنی داری برای هر ترم، به استثنای **STYLE**، کوچک تر از **0.05** است. بنابراین هر ترم، به استثنای **STYLE**، از نظر آماری معنی داری هستند **(b)**.

Dependent Variable: Amount spent

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	468402.996 <sup>a</sup>	5	93880.599	11.092	.000	.138
Intercept	39359636.4	1	39359636	4650.3	.000	.931
GENDER	158037.442	1	158037.44	18.672	.000	.051
STYLE	33506.210	2	16753.105	1.979	.140	.011
GENDER * STYLE	69858.325	2	34929.163	4.127	.017	.023
Error	2920058.824	345	8463.939			
Total	59475118.4	351				
Corrected Total	3399461.820	350				

<sup>a</sup> R Squared = .138 (Adjusted R Squared = .123)

<sup>b</sup> . The Sig. Level is .05. The Level is .05.



ادامه فصل را در مرجع کاربردی **SPSS 20 (32)** دنبال نمایید.



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