

# فصل هشتم

## توزیع مشترک دو صفت متغیر

هدفهای رفتاری: در پایان این فصل، فراگیر باید بتواند:

- ۱- جامعه را برحسب دو صفت متغیر کمی گروه‌بندی نماید.
- ۲- جدول توزیع فراوانی‌های دو صفت متغیر را تشکیل دهد.
- ۳- توزیعهای حاشیه‌ای  $x$  و  $y$  را به‌دست آورد.
- ۴- توزیعهای شرطی  $y$  برحسب  $x$  و  $x$  برحسب  $y$  را به‌دست آورد.
- ۵- مشخصه‌های عددی برای دو صفت متغیر را به تفکیک محاسبه نماید.
- ۶- مفهوم پراکندگی توأم دو صفت متغیر را بیان نماید.

در فصلهای گذشته، راجع به مطالعه جامعه‌ها با یک صفت متغیر بحث کردیم. ولی اغلب برای یک محقق نیاز می‌شود که اعضاء جامعه را توسط چند صفت متغیر با هم مورد مطالعه قرار دهد. در این فصل فقط به مطالعه دو صفت متغیر از اعضاء جامعه می‌پردازیم.

### توزیع مشترک دو صفت متغیر $x$ و $y$

اگر برای اعضاء جامعه دو صفت متغیر  $x$  و  $y$  با هم اندازه‌گیری شوند، آنگاه برای هریک از اعضاء یک زوج اندازه یا مقدار به‌دست خواهد آمد. مجموع زوجهای به‌دست آمده از نتایج مشاهدات مقادیر دو صفت را می‌توان به‌صورت زیر نشان داد.

شماره اعضاء	۱	۲	۳	$n$
اندازه صفت $x$	$x_1$	$x_2$	$x_3$	$x_n$
اندازه صفت $y$	$y_1$	$y_2$	$y_3$	$y_n$

چنانچه حجم جامعه بزرگ باشد، به‌طوری که هر زوج از اندازه‌های  $x$  و  $y$  بیش از یک بار

تکرار شوند، لزوماً نتایج مشاهدات را در جدول توزیع فراوانی دو بعدی وارد می‌کنند که در آن فراوانیها (F) با دو اندیس i برای صفت x و j برای صفت y مشخص می‌گردند، یعنی به صورت  $F_{ij}$  نشان داده می‌شوند. مجموعه دو صفت متغیر را دستگاه دو متغیر نیز می‌نامند. جدول ۱ توزیع فراوانیها را برای دو صفت متغیر x و y نشان می‌دهد:

جدول ۱

x \ y	y						$F_{i\cdot}$
	$y_1$	$y_2$	.....	$y_j$	.....	$y_t$	
$x_1$	$F_{11}$	$F_{12}$	.....	$F_{1j}$	.....	$F_{1t}$	$F_{1\cdot}$
$x_2$	$F_{21}$	$F_{22}$	.....	$F_{2j}$	.....	$F_{2t}$	$F_{2\cdot}$
$x_i$	$F_{i1}$	$F_{i2}$	.....	$F_{ij}$	.....	$F_{it}$	$F_{i\cdot}$
$x_s$	$F_{s1}$	$F_{s2}$	.....	$F_{sj}$	.....	$F_{st}$	$F_{s\cdot}$
$F_{\cdot j}$	$F_{\cdot 1}$	$F_{\cdot 2}$	.....	$F_{\cdot j}$	.....	$F_{\cdot t}$	n

در این جدول اگر فراوانیهای درون جدول ( $F_{ij}$ ) را روی سطر جمع کنیم خواهیم داشت:

$$\sum_{j=1}^t F_{ij} \cdot F_{i1} \cdot F_{i2} \cdot \dots \cdot F_{it} \cdot F_{i\cdot}$$

چون فراوانیها روی اندیس j جمع شده‌اند بنابراین جواب حاصل به صورت  $F_{i\cdot}$  نوشته خواهد شد. هم‌چنین اگر فراوانیها را روی اندیس i جمع بیندیم جواب حاصل به صورت  $F_{\cdot j}$  نوشته خواهد شد.

$$\sum_{i=1}^s F_{ij} \cdot F_{1j} \cdot F_{2j} \cdot \dots \cdot F_{sj} \cdot F_{\cdot j}$$

بنابراین جمع سطرهای جدول به صورت  $F_{1\cdot}, F_{2\cdot}, \dots, F_{s\cdot}$  در خواهد آمد و اگر آنها را روی اندیس i جمع کنیم حاصل حجم جامعه یعنی n را خواهد داد. و اگر  $F_{\cdot j}$  ها را روی j جمع کنیم نتیجه برابر حجم جامعه یعنی n خواهد شد.

$$\begin{matrix} s \\ i, 1 \end{matrix} F_{i0} \cdot F_{i1} \cdot F_{i2} \cdot \dots \cdot F_{is} \cdot n$$

$$\begin{matrix} t \\ j, 1 \end{matrix} F_{j0} \cdot F_{j1} \cdot F_{j2} \cdot \dots \cdot F_{jt} \cdot n$$

مثال ۱: فرض کنید از جامعه دانش‌آموزان یک دبیرستان دو صفت متغیر  $x$  و  $y$ ، که در آن  $x$  سن دانش‌آموزان و  $y$  تعداد افراد خانوار دانش‌آموزان، پرسش گردیده نتایج به صورت جدول توزیع فراوانیها (جدول ۲) به دست آمده باشد.

جدول ۲

جمع سطر	تعداد افراد خانوار $y$					
	۲	۳	۴	۵	۶	
سن $x$						
۱۴	۳	۴	۶	۱۲	۵	۳۰
۱۵	۲	۷	۱۰	۱۱	۵	۳۵
۱۶	۴	۵	۹	۱۰	۴	۳۲
۱۷	۶	۴	۱۲	۱۰	۵	۳۷
۱۸	۳	۳	۷	۸	۷	۲۸
۱۹	۲	۶	۵	۴	۱	۱۸
جمع ستون	۲۰	۲۹	۴۹	۵۵	۲۷	۱۸۰

اعداد داخل جدول نشانه فراوانیهای هر زوج از دو صفت متغیر می‌باشد. برای مثال ۴ دانش‌آموز ۱۴ ساله تعداد نفرات خانوارشان ۳ نفره است و هم چنین ۱۰ دانش‌آموز ۱۵ ساله تعداد نفرات خانوارشان ۴ نفره می‌باشد.

### توزیعهای حاشیه‌ای

از این جدول دوبعدی می‌توان توزیع هر یک از صفات متغیر را جداگانه به دست آورد. مثلاً بدون در نظر گرفتن صفت  $y$ ، می‌توان توزیع صفت متغیر  $x$  را به دست آورد.

جدول ۳

x	$x_1$	$x_2$	.....	$x_i$	.....	$x_s$	جمع
فراوانی $F_{i.}$	$F_{1.}$	$F_{2.}$	.....	$F_{i.}$	.....	$F_{s.}$	n

برای مثال ۱ توزیع سن دانش‌آموزان را بدون در نظر گرفتن تعداد افراد خانوار تشکیل می‌دهیم :

جدول ۴

x سن	۱۴	۱۵	۱۶	۱۷	۱۸	۱۹	جمع
فراوانی $F_{i.}$	۳۰	۳۵	۳۲	۳۷	۲۸	۱۸	۱۸۰

چون فراوانیها برای مقادیر صفت x (سن) از ستون کناری جدول قرار داده شده، به چنین توزیعی، توزیع کناری یا توزیع حاشیه‌ای x گویند. این توزیع نشان می‌دهد که توزیع سن دانش‌آموزان این دبیرستان چگونه است. همین‌طور می‌توان توزیع صفت y را به‌طور جداگانه از روی جدول دوبعدی ۱ به‌دست آورد :

جدول ۵

y	$y_1$	$y_2$	.....	$y_j$	.....	$y_t$	جمع
فراوانی $F_{.j}$	$F_{.1}$	$F_{.2}$	.....	$F_{.j}$	.....	$F_{.t}$	n

چنین توزیعی را توزیع حاشیه‌ای y نامند. چون در برابر مقادیر صفت y فراوانیها را از سطر پایین جدول ۱ قرار داده‌ایم.

توزیع حاشیه‌ای y را برای مثال ۱ به‌صورت جدول (۶) زیر می‌نویسیم :

جدول ۶

y (تعداد افراد خانوار)	۲	۳	۴	۵	۶	جمع
فراوانی $F_{.j}$	۲۰	۲۹	۴۹	۵۵	۲۷	۱۸۰

این توزیع به ما نشان می‌دهد که دانش‌آموزان این دبیرستان از نظر تعداد افراد خانوارشان چگونه توزیع شده‌اند.

### توزیعهای شرطی

هرگاه بخواهیم برای اندازه معینی از صفت  $y$ ، توزیع صفت  $x$  را بدانیم و یا برعکس برای اندازه یک  $x$  معین، توزیع صفت  $y$  را به دست آوریم. می‌توان با قرار دادن مقادیر صفت  $x$  در جدول و فراوانیهای متناظر با  $y$  معین را در برابر آنها، توزیع صفت  $x$  برحسب  $y$  معین را به دست آورد و آن را به صورت  $X|Y$  نوشت و برای توزیع  $y$  برحسب  $x$  معین آن را به صورت  $Y|X$  می‌نویسیم. به چنین توزیعهایی، توزیع شرطی  $x$  برحسب  $y$  و توزیع شرطی  $y$  برحسب  $x$  گویند. مثلاً اگر بخواهیم توزیع سن دانش‌آموزان را برای خانوارهای ۴ نفره به دست آوریم (با توجه به جدول ۲) خواهیم داشت:

جدول ۷

$X y$ . ۴	۱۴	۱۵	۱۶	۱۷	۱۸	۱۹	جمع
فراوانی $F_{i4}$	۶	۱۰	۹	۱۲	۷	۵	۴۹

ملاحظه می‌شود که فراوانیها از ستون ۴.  $y$  در برابر مقادیر صفت  $x$  نوشته شده است. هم‌چنین اگر توزیع نفرات خانوار برای دانش‌آموزان ۱۷ ساله موردنظر باشد، می‌توان آن را به صورت زیر به دست آورد:

جدول ۸

$Y x$ . ۱۷	۲	۳	۴	۵	۶	جمع
فراوانی $F_{17j}$	۶	۴	۱۲	۱۰	۵	۲۷

در اینجا نیز فراوانیها از سطر ۱۷.  $x$  در برابر مقادیر  $y$  نوشته شده است. از توزیعهای شرطی برای پی بردن به وجود بستگی بین دو صفت متغیر  $x$  و  $y$  استفاده می‌شود.

## مشخصه‌های عددی برای توزیع مشترک دو صفت متغیر x و y

برای محاسبه میانگین صفت متغیر x و میانگین صفت متغیر y از توزیعهای حاشیه‌ای x و y استفاده نموده و مطابق طریقه محاسبه میانگین حسابی، می‌توان آنها را به دست آورد:

فرمولهای محاسباتی میانگینها به قرار زیر است:

$$\bar{X} = \frac{\sum_{i=1}^s F_{i\cdot} X_i}{n} \quad (1)$$

$$\bar{Y} = \frac{\sum_{j=1}^t F_{\cdot j} Y_j}{n} \quad (2)$$

برای مثال ۱، میانگین صفت متغیر x (سن دانش‌آموزان) و میانگین صفت متغیر y (تعداد افراد خانوار) را محاسبه می‌کنیم:

$$\bar{X} = \frac{\sum_{i=1}^s F_{i\cdot} X_i}{n} = \frac{30 \cdot 14 + 35 \cdot 15 + 32 \cdot 16 + 37 \cdot 17 + 28 \cdot 18 + 18 \cdot 19}{180}$$

$$= \frac{2932}{180} = 16 \frac{29}{29}$$

به عبارت بهتر میانگین سن دانش‌آموزان ۱۶/۲۹ سال می‌باشد.

$$\bar{Y} = \frac{\sum_{j=1}^t F_{\cdot j} Y_j}{n} = \frac{20 \cdot 2 + 29 \cdot 3 + 49 \cdot 4 + 55 \cdot 5 + 27 \cdot 6}{180}$$

$$= \frac{760}{180} = 4 \frac{22}{22}$$

یعنی متوسط تعداد افراد خانوار دانش‌آموزان ۴/۲۲ نفر می‌باشد.

هم‌چنین می‌توان پراکندگی (واریانس) هر یک از صفات متغیر x و y را توسط روابط زیر محاسبه نمود:

$$V(X) = \frac{\sum_{i=1}^s F_{i\cdot} X_i^2}{n} - \bar{X}^2 \quad (3)$$

$$V(Y) = \frac{\sum_{j=1}^t F_{\cdot j} Y_j^2}{n} - \bar{Y}^2 \quad (4)$$

برای مثال ۱ واریانس صفت متغیر X و واریانس صفت متغیر Y را محاسبه می‌کنیم:

$$V(X) = \frac{\sum_{i=1}^s F_i X_i^2}{n} - \bar{X}^2$$

$$= \frac{30 \cdot 14^2 + 35 \cdot 15^2 + 32 \cdot 16^2 + 37 \cdot 17^2 + 28 \cdot 18^2 + 18 \cdot 19^2}{180} - (16/29)^2$$

$$= \frac{48210}{180} - 265/36 = 267/18 - 265/36 = 2/47$$

$$V(Y) = \frac{\sum_{j=1}^t F_j Y_j^2}{n} - \bar{Y}^2$$

$$= \frac{20 \cdot 2^2 + 29 \cdot 3^2 + 49 \cdot 4^2 + 55 \cdot 5^2 + 27 \cdot 6^2}{180} - (4/22)^2$$

$$= \frac{3472}{180} - 17/81 = 19/29 - 17/81 = 1/48$$

علاوه بر پراکندگی صفات متغیر X و Y، پراکندگی توأم دو صفت متغیر X و Y را نیز محاسبه می‌کنند که آنرا کواریانس X و Y نامند و به صورت Cov(x,y) نوشته می‌شود. فرمول محاسبه کواریانس X و Y به قرار زیر است:

$$\text{Cov}(x,y) = \frac{\sum_{i,j} F_{ij} X_i Y_j}{n} - \bar{X} \bar{Y} \quad (5)$$

از کواریانس X و Y برای آگاهی از همبستگی بین دو صفت X و Y استفاده می‌کنند. هرگاه مقدار کواریانس X و Y برابر صفر شود یعنی بین دو صفت X و Y همبستگی خطی وجود ندارد و هرگاه مقدار کواریانس X و Y بزرگتر از صفر (مثبت) باشد یعنی بین دو صفت X و Y همبستگی به‌طور مستقیم وجود دارد. و چنانچه کوچکتر از صفر (منفی) باشد یعنی بین دو صفت X و Y همبستگی به‌طور معکوس وجود دارد.

برای مثال ۱ کواریانس X و Y را محاسبه می‌کنیم:

$$\text{Cov}(x,y) = \frac{\sum_{i,j} F_{ij} X_i Y_j}{n} - \bar{X} \bar{Y}$$

$$= \frac{3 \cdot 14 \cdot 2 + 4 \cdot 14 \cdot 3 + 6 \cdot 14 \cdot 4 + 12 \cdot 14 \cdot 5 + 5 \cdot 14 \cdot 6 + 2 \cdot 15 \cdot 2}{180} - \bar{X} \bar{Y}$$

$\underline{7. 15. 3. 10. 15. 4. 11. 15. 5. 5. 11. 6. 4. 16. 2. 5. 16. 3.}$   
 $\underline{9. 16. 4. 10. 16. 5. 4. 16. 6. 6. 17. 2. 4. 17. 3. 12. 17. 4.}$   
 $\underline{10. 17. 5. 5. 17. 6. 3. 18. 2. 3. 18. 3. 7. 18. 4. 8. 18. 5.}$   
 $\underline{. 7. 18. 6. 2. 19. 2. 6. 19. 3. 5. 19. 4. 4. 19. 5. 1. 19. 6.}$   
 $180$

$$\frac{12352}{180} \cdot 16/29. 4/22. 68/62. 68/74 \dots 0/12$$

برای سهولت محاسبات می‌توان کلیه محاسبات را توسط جدول زیر انجام داد: (جدول ۹)

جدول ۹

$\begin{matrix} y \\ x \end{matrix}$	۲	۳	۴	۵	۶	$F_{i.}$	$F_{i.}X_i$	$F_{i.}X_i^2$	$F_{.j}Y_j$	$X_i \cdot F_{ij}Y_j$
۱۴	۳	۴	۶	۱۲	۵	۳۰	۴۲۰	۵۸۸۰	۱۳۲	۱۸۴۸
۱۵	۲	۷	۱۰	۱۱	۵	۳۵	۵۲۵	۷۸۷۵	۱۵۰	۲۲۵۰
۱۶	۴	۵	۹	۱۰	۴	۳۲	۵۱۲	۸۱۹۲	۱۳۳	۲۱۲۸
۱۷	۶	۴	۱۲	۱۰	۵	۳۷	۶۲۹	۱۰۶۹۳	۱۵۲	۲۵۸۴
۱۸	۳	۳	۷	۸	۷	۲۸	۵۰۴	۹۰۷۲	۱۲۵	۲۲۵۰
۱۹	۲	۶	۵	۴	۱	۱۸	۳۲۲	۶۴۹۸	۶۸	۱۲۹۲
$F_{.j}$	۲۰	۲۹	۴۹	۵۵	۲۷	۱۸۰	۲۹۳۲	۴۸۲۱۰	۷۶۰	۱۲۳۵۲
$F_{.j}Y_j$	۴۰	۸۷	۱۹۶	۲۷۵	۱۶۲	۷۶۰	..	....	.	
$F_{.j}Y_j^2$	۸۰	۲۶۱	۷۸۴	۱۳۷۵	۹۷۲	۳۴۷۲				

محاسبه اندازه همبستگی بین دو صفت متغیر  $x$  و  $y$  و تفسیر آنرا در جلد دوم کتاب روشهای

آماري از نظر خواهید گذراند.



## سؤالها و تمرینها ؟

- ۱- توزیع مشترک دو صفت متغیر چیست؟
- ۲- هدف از مطالعه جامعه‌ها با بیش از یک صفت متغیر چیست؟
- ۳- فراوانی گروه  $z$  نام چیست؟
- ۴- توزیع حاشیه‌ای صفت متغیر  $x$  چگونه به دست می‌آید؟
- ۵- توزیع حاشیه‌ای صفت متغیر  $y$  چگونه به دست می‌آید؟
- ۶- توزیع شرطی صفت متغیر  $x$  بر حسب  $y$  چه مفهومی دارد؟
- ۷- توزیع شرطی صفت متغیر  $y$  بر حسب  $x$  چه مفهومی دارد؟
- ۸- از توزیع‌های شرطی برای چه منظوری استفاده می‌شود؟
- ۹- در توزیع مشترک دو صفت متغیر، میانگین حسابی دو صفت چگونه محاسبه می‌شود؟
- ۱۰- در توزیع مشترک دو صفت متغیر، پراکندگی صفات چگونه محاسبه می‌شود؟
- ۱۱- کواریانس دو صفت متغیر چیست؟ فرمول آن را بنویسید.
- ۱۲- از کواریانس برای آگاهی از چه خصوصیتی از دو صفت متغیر استفاده می‌کنند؟
- ۱۳- هرگاه کواریانس دو صفت متغیر برابر صفر گردد چه مفهومی دارد؟
- ۱۴- نتایج مشاهدات دو صفت  $x$  و  $y$  در جدول زیر آمده است :

$x$	۸	۱۰	۱۱	۱۲	۱۴
$y$	۱۰	۶	۷	۳	۵

- میانگین  $x$  و میانگین  $y$  را محاسبه کنید.  
 واریانس  $x$  و واریانس  $y$  را محاسبه کنید.  
 کواریانس  $x$  و  $y$  را محاسبه کنید.

۱۵- توزیع مشترک دو صفت متغیر  $x$  و  $y$  به شرح جدول زیر است :

$y \backslash x$	۳	۴	۵	۶	$F_{i.}$
۵	۲	۳	۳	۲	۱۰
۱۰	۳	۱۰	۱۰	۲	۲۵
۱۵	۱	۲	۷	۵	۱۵
$F_{.j}$	۶	۱۵	۲۰	۹	۵۰

مطلوبست :

الف - توزیع حاشیه‌ای  $x$  و توزیع حاشیه‌ای  $y$

ب - توزیع شرطی  $y$  بر حسب  $x = ۱۰$

ج - توزیع شرطی  $x$  بر حسب  $y = ۵$

د - میانگین  $x$  و میانگین  $y$

هـ - واریانس  $x$  و واریانس  $y$

و - کواریانس  $x$  و  $y$  و تفسیر آن

## منابع و مأخذ

- ۱- مفاهیم اساسی آمار تألیف علی مدنی - انتشارات فروردین.
- ۲- آنالیز آماری (جلد اول) تألیف علی مدنی.
- ۳- آمار توصیفی تألیف ژرار کالت - ترجمه حسن صادقی از انتشارات دانشگاه فردوسی مشهد.
- ۴- آمار و کاربرد آن در مدیریت تألیف مهدی صفاری - نشر دانا.
- ۵- آمار مقدماتی (جلد اول) تألیف وونا کات - ترجمه دکتر مشکانی.
- ۶- Statistical Methods Snedecor & Cochran.
- ۷- Statistical Analysis for Managerial Decisions Boot and Cox.
- ۸- Introduction to Mathematical Statistics Paul . G . Hoel.

