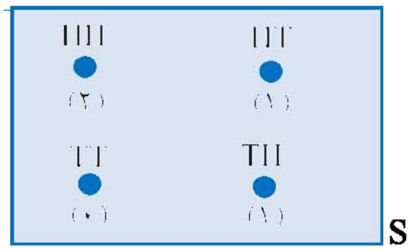
فصل ۳ متغیرهای تصادفی گسسته

۳-۱ مقدمه

در فصل قبل از اصول احتمالات به عنوان پایه استنباط آماری صحبت شد. در این فصل انواع متغیرهای تصادفی گسسته، توزیع‌های احتمالی و سایر مفاهیم مرتبط با آنها بررسی می‌شوند. همچنین از بین توزیع‌های رایج گسسته، توزیع‌های برنولی، دوجمله‌ای و پواسن معرفی می‌گردد.

۳-۲ متغیر تصادفی

متغیر تصادفی تابعی است که یک عدد حقیقی را به هر عضو فضای نمونه نسبت می‌دهد. به عنوان مثال نمودار ون زیر را در نظر بگیرید که فضای نمونه در پرتاب همزمان دو سکه را نشان می‌دهد. تعداد شیر در هر آزمایش که با x نشان داده می‌شود می‌تواند ۰، ۱ یا ۲ باشد. در اینجا x تابعی است که دامنه آن فضای نمونه یعنی $S=\left\{HH,HT,TH,TT\right\}$ و تکیه‌گاه یا برد آن زیرمجموعه‌ای از اعداد حقیقی R به صورت $S_x=\left\{۰,۱,۲\right\}$ است. تکیه‌گاه، مجموعه مقادیر ممکن متغیر تصادفی است و به صورت S_x نشان داده می‌شود.



نمودار ۳-۱. نمودار ون در آزمایش پرتاب دو سکه.

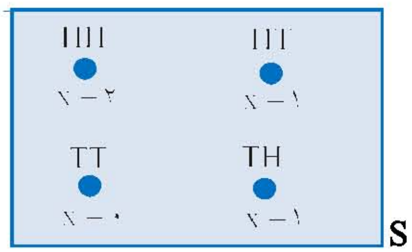
۳-۳ متغیر تصادفی گسسته یا پیوسته

هر گاه تکیه‌گاه متغیر تصادفی x یک مجموعه عددی شمارا باشد، آن را گسسته و اگر یک فاصله یا اجتماع چند فاصله باشد، آن را پیوسته گویند. تعداد شیر مشاهده شده در پرتـاب همزمـان دو سکه ($x = ۰,۱,۲$) و یا تعداد نماینده‌های خواهان استیضاح یـک وزیـر در یـک نمونـه ۵۰ نفـری ($x = ۰,۱,۲,۳,\cdots,۵۰$) نمونه‌هایی از متغیر تصادفی گسسته هستند. طول عمر یک نـوع لامـپ بـر حسب دقیقه ($۰ \leq x < \infty$) یا زمان مورد نیاز یک دانشجو برای پاسخ به یـک برگـه در حـداکثر ۹۰ دقیقه ($۰ \leq x \leq ۹۰$)، نمونه‌هایی از متغیر تصادفی پیوسته هستند. در این فصل متغیر تصادفی گسسته و در فصل چهارم متغیر تصادفی پیوسته مطرح خواهد شد.

۳-۴ تابع چگالی احتمال

تابع چگالی احتمال یا به طور خلاصه تابع احتمال یک متغیر تصادفی تـابعی اسـت کـه احتمـال مشاهده هر مقدار ممکن متغیر تصادفی را مشخص می‌کند، این تابع مـی‌توانـد بـه صـورت یـک نمودار، جدول یا یک فرمول باشد. هر تابع احتمال باید دارای دو ویژگی باشد. اول اینکه احتمال برای هر مقدار x بین صفر و یک باشد، یعنی $۰ \leq P(x) \leq ۱$ و دوم اینکه رابطـه $\sum_{x \in S_x}^{n} P(x) = ۱$ برقرار باشد.

مثال ۳-۱. اگر x تعداد شیر مشاهده شده در پرتاب همزمان دو سکه باشد جدول توزیع احتمال مربوطه را به دست آورید.



نمودار ۳-۲. نمودار ون برای آزمایش پرتاب دو سکه.

همانگونه که در نمودار نشان داده شده است متغیر تصادفی x می‌تواند هر کدام از مقـادیر ۰، ۱ و

۲ باشد. احتمال مشاهده هر کدام از اعضای نمونه به صورت زیر محاسبه می‌شود:

$$P(۰) = P(TT) = \frac{۱}{۴}$$
$$P(۱) = P(TH) + P(HT) = \frac{۱}{۴} + \frac{۱}{۴} = \frac{۱}{۲}$$
$$P(۲) = P(HH) = \frac{۱}{۴}$$

جدول ۳-۱. جدول تابع چگالی احتمال برای تعداد شیر در پرتاب دو سکه.

x	۰	۱	۲
$P(x)$	۰/۲۵	۰/۵	۰/۲۵

۳-۵ میانگین و واریانس متغیر تصادفی گسسته

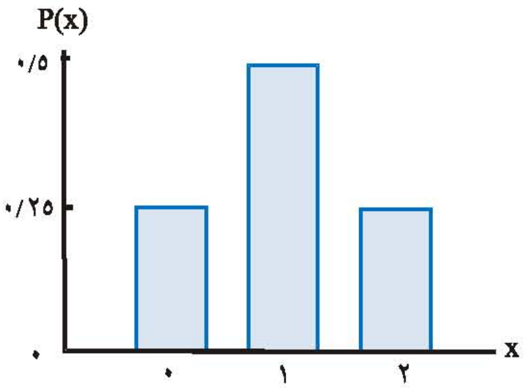
چون مقدار یک متغیر تصادفی، قبل از رخداد پدیده تصادفی مشخص نیست، از این رو دانستن متوسط مقدار آن متغیر یا میزان پراکندگی آن از اهمیت خاصی برخوردار است. کمیت‌های میانگین و واریانس به ترتیب، مقدار متوسط و میزان پراکندگی مقادیر آن نسبت به میانگین را مشخص می‌کند. مشخص است که در تعداد زیادی پرتاب همزمان دو سکه، در ۲۵٪ موارد، x یا تعداد شیر مشاهده شده برابر با صفر است، در ۵۰٪ موارد x=۱ و در ۲۵٪ موارد نیز x=۲ است. در نتیجه میانگین تعداد شیر مشاهده شده برابر است با:

$$\mu = ۰(\frac{۱}{۴}) + ۱(\frac{۱}{۲}) + ۲(\frac{۱}{۴}) = ۰ + \frac{۱}{۲} + \frac{۱}{۲} = ۱$$

یعنی اگر احتمال هر مقدار x به عنوان فراوانی نسبی رخداد آن مقدار در نظر گرفته شود، میانگین یا مقدار مورد انتظار متغیر تصادفی x برابر

$$\mu_x = E(x) = \sum x.p(x)$$

است. میانگین x را امید ریاضی x نیز می‌نامند و با $E(x)$ نشان می‌دهند. برای هر متغیر تصادفی x، میانگین، متوسط و مقدار مورد انتظار همگی یک معنی دارند.



نمودار ۳-۳. نمودار میله‌ای توزیع احتمال برای پرتاب همزمان دو سکه.

همچنین واریانس متغیر تصادفی x عبارتست از متوسط مربع فواصل xها از میانگین جامعه (μ). اگر احتمال هر مقدار x فراوانی نسبی رخداد آن مقدار را نشان دهد، می‌توان میانگین مقادیر مختلف $(x-\mu)^{۲}$ را به صورت

$$\sigma^{۲} = E(x-\mu)^{۲} = \sum (x-\mu)^{۲} p(x)$$

به دست آورد. همچنین می‌توان نشان داد که رابطه $E(x-\mu)^{۲} = E(x^{۲}) - \mu^{۲}$ برقرار می‌باشد. عبارت $E(x-\mu)^{۲}$ امید ریاضی توان دوم انحرافات از میانگین نیز نامیده می‌شود.

۳-۶ توزیع دوجمله‌ای

در خیلی از آزمایش‌ها نتیجه از دو حالت شکست یا پیروزی خارج نیست. به این آزمایش‌ها آزمایش برنولی گفته می‌شود و در آنها مقدار متغیر تصادفی برای حالت شکست، صفر و برای حالت پیروزی، یک در نظر گرفته می‌شود. احتمال پیروزی برابر p (عددی بین صفر و یک) و احتمال شکست برابر $q = ۱ - p$ است. توزیع احتمال مربوطه را می‌توان به هر یک از صورت‌های زیر نیز بیان کرد.

x	۰	۱
$P(x)$	q	p

$$P(x) = p^{x} q^{۱-x}, \qquad (x = ۰,۱)$$

حال اگر یک آزمایش برنولی n بار به طور مستقل تکرار شود (یعنی نتیجه هر آزمایش بر نتایج

آزمایش‌های دیگر تاثیری نداشته باشد) و متغیر تصادفی x با تکیه‌گاه $\{۰,۱,۲,...,n\}$ برابر مجموع پیروزی‌ها باشد در این صورت متغیر تصادفی x دارای توزیع دوجمله‌ای است و بصورت $x \sim bin(n,p)$ تعریف می‌شود.

برای مشاهده مقدار خاصی مانند x باید نتیجه x آزمایش، پیروزی و نتیجه $n-x$ آزمایش باقیمانده شکست باشد. چون آزمایش‌ها از هم مستقل‌اند، احتمال مشاهده چنین نتیجه‌ای برابر حاصلضرب احتمال‌ها یعنی $p^x q^{n-x}$ است. از طرفی، تعداد $\binom{n}{x}$ حالت مختلف برای داشتن x پیروزی و $n-x$ شکست در یک آزمایش دوجمله‌ای وجود دارد. برای مثال در ۴ بار پرتاب سکه‌ای همانگونه که در ستون دوم جدول ۳-۲ نشان داده شده است ۴ حالت مختلف برای مشاهده یک شیر $(x=۱)$ وجود دارد. یک حالت این است که سه پرتاب اول خط و پرتاب آخر شیر آید یعنی $TTTH$. حالت دوم این است که پرتاب سوم شیر آید و بقیه پرتاب‌ها خط یعنی $TTHT$. دو حالت دیگر را می‌توان به صورت $THTT$ و $HTTT$ نشان داد.

جدول ۳-۲. نتایج ممکن در پرتاب همزمان ۴ سکه با در نظر گرفتن تعداد شیر (x)به عنوان متغیر تصادفی.

$x=۰$	$x=۱$	$x=۲$	$x=۳$	$x=۴$
$TTTT$	$TTTH$	$TTHH$	$HHHT$	$HHHH$
	$TTHT$	$THHT$	$HHTH$	
	$THTT$	$HHTT$	$HTHH$	
	$HTTT$	$THTH$	$THHH$	
		$HTTH$		
		$HTHT$		
$P(۰)=\frac{۱}{۱۶}$	$P(۱)=\frac{۴}{۱۶}$	$P(۲)=\frac{۶}{۱۶}$	$P(۳)=\frac{۴}{۱۶}$	$P(۳)=\frac{۱}{۱۶}$

با توجه به مستقل بودن آزمایش‌ها، احتمال مشاهده یک شیر یعنی $P(x=۱)$ را می‌توان از جمع احتمال حالات مختلف مورد نظر به دست آورد:

$$P(\text{مشاهده یک شیر}) = P(\{HTTT\} \text{ یا } \cdots \text{ یا } \{TTTH\})$$
$$= pqqq + qpqq + qqpq + qqqp = pq^۳ + pq^۳ + pq^۳ + pq^۳$$
$$= (\frac{۱}{۲})(\frac{۱}{۲})^۳ + (\frac{۱}{۲})(\frac{۱}{۲})^۳ + (\frac{۱}{۲})(\frac{۱}{۲})^۳ + (\frac{۱}{۲})(\frac{۱}{۲})^۳ = ۴(\frac{۱}{۲})(\frac{۱}{۲})^۳ = \frac{۴}{۱۶} = ۰/۲۵$$

به طور خلاصه برای محاسبه احتمال یک پیروزی یعنی $P(x=۱)$ می‌توان نوشت

$$P(x=۱) = \binom{۴}{x}(۰/۵)^x(۰/۵)^{۴-x} = \binom{۴}{۱}(۰/۵)^۱(۰/۵)^۳ = ۰/۲۵$$

که در آن عبارت $\binom{۴}{x}$ تعداد حالات ممکن با x موفقیت یا همان تعداد ترکیبات xتایی از ۴ عضو است. عبارت $(۰/۵)^x(۰/۵)^{۴-x}$ نیز احتمال هر یک از اعضای نمونه با x موفقیت می‌باشد. احتمال سایر xها در چنین آزمایشی به همراه احتمالات مربوطه در جدول بالا آمده است که در آن x تعداد موفقیت‌ها (شیرها) می‌باشد.

به طور کلی در n بار تکرار مستقل یک آزمایش برنولی، احتمال x بار موفقیت برابر است با:

$$P(x) = \binom{n}{x} p^x (۱-p)^{n-x} \qquad Sx = \{۰,۱,۲,\cdots,n\}$$

$p =$ احتمال موفقیت در یک تکرار

$q = ۱ - p$: احتمال شکست در هر تکرار

$n =$ تعداد تکرار

$x =$ تعداد موفقیت در n تکرار

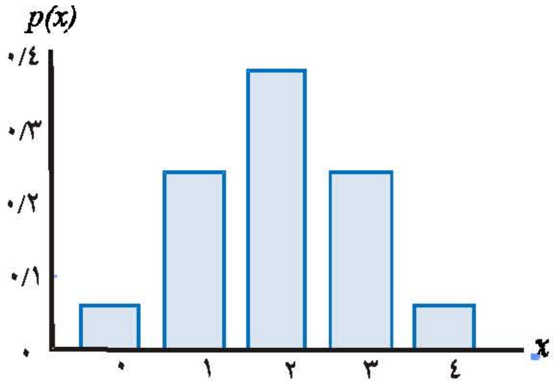
$$\binom{n}{x} = \frac{n!}{(n-x)!\,x!}$$

با فرمول بالا بدون نیاز به فهرست کردن همه اعضای نمونه در یک جدول، می‌توان به راحتی احتمالات مختلف را در توزیع دوجمله‌ای محاسبه نمود به ویژه اینکه تعداد حالات مختلف با افزایش n (تعداد تکرارهای تشکیل دهنده یک توزیع دوجمله‌ای) به شدت افزایش می‌یابد.

مثال ۳-۲. احتمال وقوع ۲ شیر در پرتاب همزمان ۴ سکه سالم برابر است با $۶p^۲q^۲$ و چون

$p = q = ۰/۵$ پس از جایگذاری، مقدار $۶\times ۰/۵^{۲} \times ۰/۵^{۲} = ۰/۳۷۵$ به دست می‌آید، و به همین ترتیب احتمال وقوع ۳ شیر در پرتاب همزمان ۴ سکه برابر است با:

$$P(x=۳)=\binom{۴}{۳}p^{۳}q=۴\times ۰/۵^{۳}\times ۰/۵=۰/۲۵$$

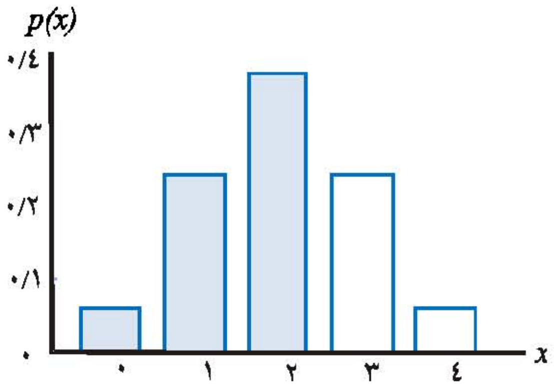


نمودار ۳-۴. نمودار توزیع احتمال متغیر تصادفی تعداد شیر در پرتاب همزمان ۴ سکه.

مثال ۳-۳. احتمال وقوع حداکثر ۲ شیر در پرتاب همزمان ۴ سکه برابر است با

$$P(x\leq ۲)=P(x=۰)+P(x=۱)+P(x=۲)=۰/۶۸۷۵$$

این احتمال برابر است با مجموع مساحت مستطیل‌های سایه‌دار در نمودار ۳-۵. نمودار مربوط به توزیع دوجمله‌ای فقط برای حالت $p = q = ۰/۵$ متقارن است. در غیر اینصورت توزیع نامتقارن می‌شود.



نمودار ۳-۵. احتمال وقوع حداکثر ۲ شیر در پرتاب همزمان ۴ سکه.

جدول ۳-۳. محاسبه احتمال در توزیع دوجمله‌ای با استفاده از نرم‌افزار Minitab.

در n بار تکرار یک آزمایش برنولی اگر p احتمال موفقیت در هر تکرار باشد P(x) در مسیر زیر محاسبه می‌شود:

Calc > Probability Distributions > Binomial…

مثال: احتمال ۵ موفقیت در ۸ بار تکرار یک آزمایش برنولی در صورتی که p برابر با ۰/۴ باشد چقدر است؟

حل:

۱. گزینه probability را انتخاب کنید.

۲. در مقابل Number of trials عدد ۸ و در مقابل Event probability عدد ۰/۴ را وارد کنید.

۳. گزینه Input constant را انتخاب و در مقابل آن عدد ۵ را وارد کنید. سپس بر روی *OK* کلیک کنید.

```
Binomial with n = 8 and p = 0.4
x   P( X = x )
5    0.123863
```

مثال: احتمال حداکثر ۵ موفقیت را در ۸ بار تکرار یک آزمایش برنولی در صورتی که p برابر با ۰/۴ باشد به دست آورید.

۱. گزینه Cumulative probability را انتخاب کنید.

۳. در مقابل Number of trials عدد ۸ و در مقابل Event probability عدد ۰/۴ را وارد کنید.

۴. گزینه Input constant را انتخاب و در مقابل آن عدد ۵ را وارد کنید. سپس بر روی OK کلیک کنید.

```
Binomial with n = 8 and p = 0.4

x  P( X <= x )
5     0.950193
```

به طور کلی در n بار تکرار یک آزمایش برنولی، جملات متوالی بسط عبارت $(p+q)^n$، احتمال‌های مربوط به ۰، ۱، ۲، ...، n موفقیت در n تکرار مذکور است که به طور مستقل از هم انجام شده‌اند. برای مثال در ٤ بار پرتاب یک سکه می‌توان نوشت

$$(p+q)^٤ = \binom{٤}{٠} q^٤ + \binom{٤}{١} q^٣ p + \binom{٤}{٢} p^٢ q^٢ + \binom{٤}{٣} qp^٣ + \binom{٤}{٤} p^٤$$

لذا احتمال ۰، ۱، ۲، ۳ و ٤ بار مشاهده شیر در این آزمایش دوجمله‌ای به ترتیب برابر با $q^٤$، $٤q^٣p$، $٦p^٢q^٢$، $٤qp^٣$ و $p^٤$ یعنی ۰/۰٦۲٥، ۰/۲٥، ۰/۳۷٥، ۰/۲٥ و ۰/۰٦۲٥ خواهد بود.

یک راه ساده برای بسط دادن عبارت $(p+q)^n$ بدین ترتیب است که ابتدا جملات به صورت

p^n، $p^{n-1}q^1$، $p^{n-2}q^2$، ...، q^n نوشته می‌شوند. ضریب ۱ را برای جمله اول قرار داده و برای جملات بعدی، ضریب جمله قبلی در توان q ضرب شده و بر تعداد جملات قبلی تقسیم می‌شود. برای مثال ضریب جمله سوم در بسط دو جمله‌ای پایین برابر است با ضریب جمله قبلی یعنی ٦ ضرب در ٥ (توان q در جمله قبلی) تقسیم بر ۲ (تعداد جملات قبلی) که برابر با ۱۵ است.

$$(p+q)^6 = q^6 + 6q^5p + 15q^4p^2 + 20q^3p^3 + 15q^2p^4 + 6qp^5 + p^6$$

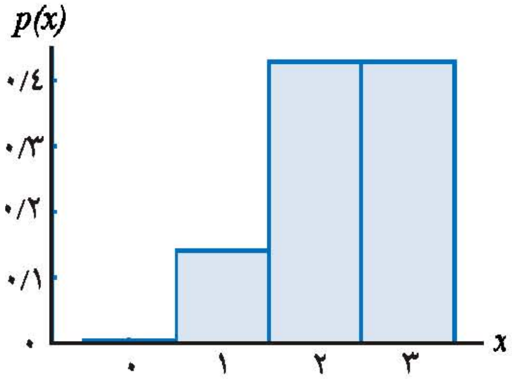
مثال ۳-٤. فرض کنید که در جمعیت F_2حاصل از تلاقی دو واریته، دو نوع بذر سیاه و سفید به نسبت ۳ به ۱ وجود دارد. در اینجا جمعیت F_2 به دو پیشامد سیاه و سفید تقسیم می‌شود. اگر نمونه‌های سه تایی از این جمعیت انتخاب شود، در هر نمونه ممکن است ۰، ۱، ۲ یا ۳ بذر سیاه وجود داشته باشد. در این مثال احتمال مشاهده ۰، ۱، ۲ یا ۳ بذر سیاه در یک نمونه ۳ تایی به صورت زیر محاسبه می‌شود:

$$P(x=0) = \binom{3}{0} p^0 q^{3-0} = 1(\tfrac{3}{4})^0(\tfrac{1}{4})^{3-0} = \frac{1}{64}$$
$$P(x=1) = \binom{3}{1} p^1 q^{3-1} = 3(\tfrac{3}{4})^1(\tfrac{1}{4})^{3-1} = \frac{9}{64}$$
$$P(x=2) = \binom{3}{2} p^2 q^{3-2} = 3(\tfrac{3}{4})^2(\tfrac{1}{4})^{3-2} = \frac{27}{64}$$
$$P(x=3) = \binom{3}{3} p^3 q^{3-3} = 1(\tfrac{3}{4})^3(\tfrac{1}{4})^{3-3} = \frac{27}{64}$$

جدول ۳-٤. جدول توزیع احتمال برای آزمایش دوجمله‌ای مشاهده تعداد بذور سیاه (x) در یک نمونه سه‌تایی بذر F_2.

x	۰	۱	۲	۳
$P(x)$	۰/۰۱۵٦۲٥	۰/۱٤۰٦۲٥	۰/٤۲۱۸۷٥	۰/٤۲۱۸۷٥

با توجه به اینکه p و q با هم برابر نیستند، پس شکل توزیع نامتقارن و به صورت زیر است.



نمودار ۳-٦. نمودار توزیع دوجمله‌ای برای تعداد بذور سیاه در نمونه‌های سه تایی بذر از جمعیت F_2 حاصل از تلاقی دو واریته متفاوت.

مثال ۳-۵. داروی خاصی بر روی ۷۰٪ از افراد مبتلا به یک نوع بیماری مؤثر است. فرض کنید ۵ بیمار تحت تأثیر دارو قرار گرفته و x تعداد افراد بهبود یافته از بین این ۵ نفر باشد. توزیع احتمال x در زیر آورده شده است. میانگین (μ) و انحراف معیار (σ) را به دست آورده و نتایج را تفسیر کنید. بافت‌نگار $P(x)$ را رسم کرده و فاصله $\mu \pm ۲\sigma$ را بر روی آن مشخص کنید. با استفاده از قاعده چبیشف و قاعده تجربی (فصل اول، بند ۱-۱۲-٤) احتمال قرار گرفتن x در خارج از این محدوده را مشخص کنید. نتیجه را با احتمال واقعی مقایسه نمائید.

x	۰	۱	۲	۳	٤	۵
$P(x)$	۰/۰۰۲	۰/۰۲۹	۰/۱۳۲	۰/۳۰۹	۰/۳٦۰	۰/۱٦۸

از آنجا که میانگین یک توزیع نشان‌دهنده متوسط مقدار متغیر تصادفی است، بنابراین

$$\mu = E(x) = \sum xp(x) = ۰(۰/۰۰۲) + ۱(۰/۰۲۹) + \cdots + ۵(۰/۱٦۸) = ۳/۵$$

درنتیجه دارو از هر ۵ نفر بیمار به طور میانگین ۳/۵ نفر را بهبود می‌بخشد. واریانس x نیز به صورت

$$\sigma^۲ = E(x-\mu)^۲ = \sum (x-\mu)^۲ p(x)$$
$$= (۰ - ۳/۵)^۲(۰/۰۰۲) + (۱ - ۳/۵)^۲(۰/۰۲۹) + \cdots + (۵ - ۳/۵)^۲(۰/۱٦۸) = ۱/۰۵$$

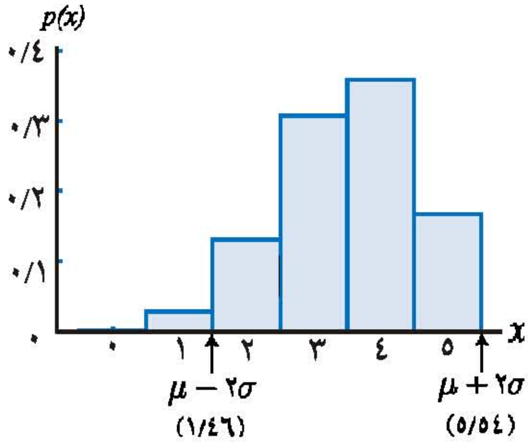
محاسبه می‌شود و از آنجا انحراف معیار برابر با $\sigma = ۱/۰۲$ است.

بافت‌نگار $P(x)$ نیز در زیر به همراه فاصله $\mu \pm ۲\sigma$ یعنی $۳/۵ \pm ۲(۱/۰۲)$ یا ۱/٤٦ تا ۵/۵٤ نشان داده شده است. میانگین (۳/۵) در مرکز این فاصله واقع شده است. بر اساس قاعده چبیشف، در هر توزیعی حداقل ۷۵٪ و بر اساس قاعده تجربی، در توزیع‌های تقریباً متقارن ۹۵٪ مقادیر xها

در فاصله دو انحراف معیار از میانگین قرار دارند. فاصله مذکور بر اساس قاعده تجربی در این مثال ١/٤٦ تا ٥/٥٤ می‌باشد. با توجه با شکل بالا، xهای برابر با ٢، ٣، ٤ و ٥ در این دامنه قرار گرفته‌اند که مجموع احتمالات آنها برابر است با:

$$P(۲) + P(۳) + P(٤) + P(٥) = ۰/۱۳۲ + ۰/۳۰۹ + ۰/۳٦۰ + ۰/۱٦۸ = ۰/۹٦۹$$

بنابراین ٩٦/٩٪ مقادیر xها در فاصله دو انحراف معیار از میانگین قرار گرفته‌اند که در توافق با قاعده چبیشف و قاعده تجربی می‌باشد.



نمودار ٣-٧. بافت‌نگار $p(x)$ برای مثال ٣-٥.

احتمال اینکه از بین ٥ نفر تیمار شده با دارو کمتر از ٢ نفر خوب شوند، در خارج از فاصله $\mu \pm ۲\sigma$ واقع شده و برابر با $P(x \leq ۱) = P(۰) + P(۱) = ۰/۰۰۲ + ۰/۰۲۹ = ۰/۰۳۱$ است. در نتیجه با وجود ٥ بیمار احتمال کمی (٠/٠٣) وجود دارد که کمتر از ٢ فرد خوب شوند. به عبارتی انتظار نمی‌رود که با این دارو کمتر از ٢ فرد بهبود یابند.

٣-٦-٢ میانگین و انحراف معیار در توزیع دو جمله‌ای

میانگین و واریانس توزیع دوجمله‌ای مانند هر توزیع گسسته دیگری از روابط $\mu = E(x) = \sum x.p(x)$ و $\sigma^۲ = E(x-\mu)^۲ = \sum (x-\mu)^۲ p(x)$ و پس از یک سری محاسبات جبری به صورت زیر بدست می‌آیند.

میانگین، واریانس و انحراف معیار در توزیع دو جمله‌ای

$$\mu = np, \quad \sigma^2 = npq, \quad \sigma = \sqrt{npq}$$

مثال ۳-۶. جدول توزیع احتمال متغیر تصادفی تعداد شیرها (x) در پرتاب همزمان ۴ سکه به صورت زیر است. میانگین و واریانس متغیر تصادفی x را به دست آورید.

x	۰	۱	۲	۳	۴
$P(x)$	۱÷۱۶	۴÷۱۶	۶÷۱۶	۴÷۱۶	۱÷۱۶

میانگین به صورت زیر محاسبه می‌شود:

$$\mu = \sum x \ p(x) = (0 \times \frac{1}{16}) + (1 \times \frac{4}{16}) + (2 \times \frac{6}{16}) + (3 \times \frac{4}{16}) + (4 \times \frac{1}{16}) = 2$$

$$\sigma^2 = \sum (x-\mu)^2 p(x)$$

$$= (0-2)^2 \frac{1}{16} + (1-2)^2 \frac{4}{16} + (2-2)^2 \frac{6}{16} + (3-2)^2 \frac{4}{16} + (4-2)^2 \frac{1}{16} = 1$$

واریانس را برای مثال بالا می‌توان به صورت $\sigma^2 = \sum x^2 p(x) - \mu^2 = 5 - 4 = 1$ نیز به دست آورد. محاسبه میانگین و واریانس با روش مخصوص توزیع دو جمله‌ای به صورت $\mu = np = 4 \times 0/5 = 2$ و $\sigma^2 = npq = 4 \times 0/5 \times 0/5 = 1$ می‌باشد.

۳-۷ توزیع چند جمله‌ای

اگر نتیجه آزمایشی، یکی از سه یا چند پیشامد ناسازگار و مستقل، هر کدام با احتمال ثابتی باشد، توزیع مربوط به آن، توزیع چندجمله‌ای است. مثلاً نتیجه مسابقه‌ای ممکن است مساوی، برد یا باخت و یا رنگ گل‌های حاصل از یک نوع بذر ممکن است سفید، صورتی یا قرمز باشد. تفاوت این توزیع با توزیع دوجمله‌ای این است که در توزیع دوجمله‌ای فقط دو پیشامد ممکن وجود دارد. اگر k تعداد پیشامدهای ممکن باشد، احتمال مشاهده ترکیب خاصی از پیشامدهای ممکن یعنی مشاهده x_1 بار پیشامد اول، x_2 بار پیشامد دوم نهایتاً x_k بار پیشامد kام) در n آزمایش از

رابطه

$$P(x_1, x_2, \cdots, x_k) = \binom{n}{x_1, x_2, \cdots, x_k} {p_1}^{x_1} {p_2}^{x_2} \cdots {p_k}^{x_k}$$

به دست می‌آید، که در آن $x_1, x_2, \cdots, x_k = n$ است.

مثال ۳-۷. انتظار می‌رود که گیاهان حاصل از تلاقی دو والد از لحاظ شکل ظاهری در ۴ دسته A، B، C و D به ترتیب با نسبت‌های ۹، ۳، ۳، ۱ قرار گیرند. احتمال اینکه از بین ۱۰ گیاه حاصل، ۶ گیاه در دسته A، ۲ گیاه در دسته B و یک گیاه در هر کدام از دسته‌های C و D قرار گیرد، چقدر است؟

$$P(6,2,1,1) = \binom{10}{6,2,1,1}\left(\frac{9}{16}\right)^6\left(\frac{3}{16}\right)^2\left(\frac{3}{16}\right)^1\left(\frac{1}{16}\right)^1 = 0/04$$

۳-۸ متغیر تصادفی پواسن

توزیع پواسن را توزیع پیشامدهای نادر گویند. اگر متغیر تصادفی گسسته، تعداد وقایع مورد نظر در محدوده‌ای از زمان یا مکان باشد، تحت شرایطی دارای توزیع پواسن است. متغیر تصادفی پواسن مقادیر $\{0,1,2,\ldots\}$ را اختیار می‌کند و از این رو تکیه‌گاه آن برخلاف متغیر دو جمله‌ای متناهی نیست. توزیع احتمال متغیر تصادفی پواسن به صورت زیر است.

توزیع احتمال، میانگین و واریانس برای متغیر تصادفی دارای توزیع پواسن:

$$P(x) = \frac{\lambda^x . e^{-\lambda}}{x!} \qquad S_x = \{0,1,2,\cdots\}$$

$$\mu = \sigma^2 = \lambda$$

$\lambda =$ همان پارامتر توزیع یا میانگین تعداد وقایع در یک واحد زمان، مکان یا حجم مشخص است و e برابر با ۲/۷۱۸ است.

مثال‌هایی چون تعداد مشتریان یک بانک در طول روز، تعداد تلفن‌ها به یک شرکت در یک ساعت، تعداد خرابی‌های جاده در طول یک کیلومتر، تعداد پنچری‌های تایر در یک هفته، تعداد

موارد گزارش شده از یک بیماری نادر و غیره معمولاً دارای توزیع پواسن هستند. میانگین و واریانس هر دو در توزیع پواسن برابر با پارامتر λ هستند.

مثال ۳–۸ اگر تعداد زدگی‌ها در یک متر مربع از پارچه‌ای به طور متوسط ٤ باشد. احتمال اینکه در یک متر مربع زدگی وجود نداشته باشد، چقدر است؟ احتمال ۳ زدگی چقدر است؟

$$P(x=٠)=\frac{\lambda^{x}.e^{-\lambda}}{x!}=\frac{٤^{٠}e^{-٤}}{٠!}=e^{-٤}=٠/٠١٨٣$$

$$P(x=٣)=\frac{٤^{٣}e^{-٤}}{٣!}=٠/١٩٥$$

جدول ۳–۵. محاسبه $P(x\leq ٢)$ مربوط به مثال ۳–۸ در نرم‌افزار Minitab.

مراحل:

۱. وارد مسیر زیر شوید:

Calc > Probability Distributions > Poisson…

۲. گزینه Cumulative probability (احتمال تجمعی) را انتخاب کنید.

۳. در مقابل Mean عدد ۴ و در مقابل Input constant عدد ۲ را وارد کنید.

۴. بر روی OK کلیک کنید.

```
Poisson with mean = 4

x    P( X <= x )
2       0.238103
```

چون محاسبه $\binom{n}{x}$ برای nهای بزرگ مشکل است، اگر در توزیع دوجمله‌ای n بزرگ باشد محاسبه احتمالات کار ساده‌ای نخواهد بود. اگر در توزیع دوجمله‌ای با n بزرگ، احتمال p خیلی کوچک باشد، بطوری که رابطه $np\leq ٥$ برقرار باشد، می‌توان احتمال‌ها را از طریق توزیع پواسن با تقریب خوبی محاسبه کرد. برای این کار کافی است λ برابر np اختیار شود.

مثال ۳–۹. احتمال اینکه یک کودک نسبت به یک سرم حساسیت داشته باشد ۰/۰۰۱ است. مطلوبست احتمال اینکه در بین ۲۰۰۰ کودک واکسینه شده، الف) سه کودک نسبت به سرم حساسیت داشته باشند. ب) بیش از ۲ کودک حساسیت داشته باشند.

$$\lambda = np = ۲۰۰۰ \times ۰/۰۰۱ = ۲$$

$$P(x=۳) = \frac{۲^{۳} \times e^{-۲}}{۳!} = ۰/۱۸$$

$$P(x>۲) = ۱ - P(x \le ۲) = ۱ - P(x = ۰ \text{ یا } x = ۱ \text{ یا } x = ۲)$$

$$= ۱ - (\frac{۱}{e^{۲}} + \frac{۲}{e^{۲}} + \frac{۲}{e^{۲}}) = ۱ - ۰/۶۷۶۶ = ۰/۳۲۳$$

مثال ۳-۱۰. اگر احتمال بروز یک بیماری خاص در یک توده بومی جو ۰/۰۰۱ باشد، در یک نمونه تصادفی ۵۰۰۰ بوته‌ای احتمال وقوع ۲ گیاه بیمار را محاسبه کنید. این احتمال با استفاده از توزیع دوجمله‌ای به سختی با محاسبه $\binom{۵۰۰۰}{۲}(۰/۰۰۱)^{۲}(۰/۹۹۹)^{۴۹۹۸}$ به دست آید. در حالی که با استفاده از تقریب توزیع پواسن می‌توان نوشت

$$\lambda = np = ۵۰۰۰ \times ۰/۰۰۱ = ۵$$

$$P(x=۲) = \frac{\lambda^{x}.e^{-\lambda}}{x!} = \frac{۵^{۲} \times e^{-۵}}{۲!}$$

تمرین‌ها

۳-۱. جدول توزیع احتمال زیر را در نظر بگیرید. **الف)** مقادیر μ، $\sigma^{۲}$ و σ را محاسبه کنید. **ب)** بافت‌نگار مقادیر x را رسم کرده و نقاط $\mu - ۲\sigma$ و $\mu + ۲\sigma$ را بر روی بافت‌نگار نشان دهید. **ج)** احتمال قرار گرفتن x در خارج از فاصله $\mu - ۲\sigma$ تا $\mu + ۲\sigma$ را به دست آورید.

x	۴-	۳-	۲-	۱-	۰	۱	۲	۳	۴
$P(x)$	۰/۰۲	۰/۰۷	۰/۱۰	۰/۱۵	۰/۳۰	۰/۱۸	۰/۱۰	۰/۰۶	۰/۰۲

۳-۲. جدول توزیع احتمالی متغیر تصادفی گسسته x با تکیه‌گاه $S_x = \{۱,۲,۳,۴,۵\}$ یا به عبارتی $S_x = [۱,۵]$ طراحی کرده و برای هر x به طور دلخواه یک احتمال یا $P(x)$ در نظر بگیرید. **الف)** حاصل $\sum P(x)$ را به دست آورید. **ب)** مقدار $\mu = E(x)$ و $E(x-\mu)^{۲}$ و σ را به دست آورید. **ج)** مقدار μ را در این مثال تفسیر کنید.

۳-۳. توزیع احتمال $P(x) = \binom{۶}{x}(۰/۷)^{x}(۰/۳)^{۶-x}, \ (x = ۰,۱,۲,\cdots,۶)$ را در نظر بگیرید. **الف)** متغیر تصادفی x گسسته است یا پیوسته؟ چرا؟ **ب)** نام این توزیع احتمال چیست؟ **ج)** مقادیر μ، $\sigma^{۲}$ و σ را

برای متغیر x محاسبه کنید.

۳-۴. اگر x یک متغیر تصادفی دوجمله‌ای با $n = ۵$ و $p = ۰/۵$ باشد، $P(x)$ را برای xهای ۱، ۲، ۳، ٤، و ۵ با استفاده از فرمول توزیع دوجمله‌ای به دست آورید.

۳-۵. دانشجویی در مورد ۱۰ سؤال ٤ گزینه‌ای هیچ اطلاعی ندارد. با چه احتمالی بیش از ٦ سوال از ۱۰ سؤال مذکور را درست جواب می‌دهد؟

۳-۶. اگر x متغیری تصادفی با توزیع احتمال پواسن باشد، مقادیر احتمال خواسته شده را به دست آورید:
الف) $P(x \le ۲)$ در صورتی که $\lambda = ۲$ باشد. **ب)** $P(x \le ۲)$ در صورتی که $\lambda = ۳$ باشد.

۳-۷. اگر x متغیری تصادفی دارای توزیع احتمال پواسن با $\lambda = ۱$ باشد، **الف)** بافت‌نگار $P(x)$ را برای مقادیر مختلف ۵، ٤، ۳، ۲، $x = ۱$ رسم کنید. **ب)** مقادیر μ، $\sigma^۲$ را برای متغیر x محاسبه کنید.

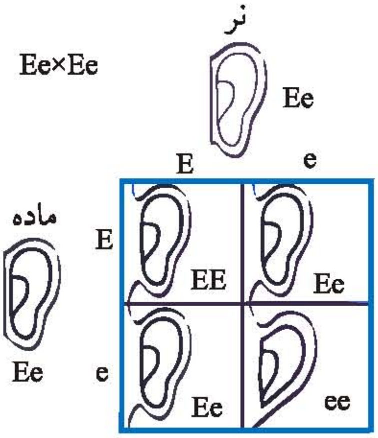
۳-۸. در بیمارستانی روزانه به طور میانگین ۲/۵ بیمار به ICU منتقل می‌شوند. اگر فقط ٤ تخت خالی در یک روز در ICU موجود باشد، با چه احتمالی بخش ICU با مشکل کمبود تخت مواجه خواهد شد؟

۳-۹. آمار نشان می‌دهد که ۰/۰۲ از مسافرین قطار بلیط خود را پس می‌دهند. اگر در یک روز ۲۵۰ نفر بلیط گرفته باشند، احتمال موارد خواسته شده را به دست آورید. **الف)** سه نفر بلیط خود را پس دهند. **ب)** بیش از سه نفر بلیط خود را پس دهند.

۳-۱۰. **الف)** با استفاده از فرمول دوجمله‌أی، عبارت $(x - y^۲)^۵$ را بسط دهید. **ب)** جمله سوم را در عبارت $(x - ۳b^۳)^۳$ بدست آورید.

۳-۱۱. در یک خانواده ٤ فرزندی مطلوب است. **الف**) احتمال وجود حداقل یک پسر ب) احتمال فقط یک پسر ج) احتمال وجود حداقل یک پسر و حداقل یک دختر.

۳-۱۲. ژنتیک‌دان‌ها از جدولی به نام جدول پانِت برای نشان دادن ترکیبات ژنتیکی ممکن حاصل از یک تلاقی استفاده می‌کنند. نمونه‌ای از این جدول که در زیر نشان داده شده است، مربوط است به فرزندان یک زوج. در انسان لاله گوش نرمال توسط آلل غالب E و لاله گوش چسبیده توسط آلل مغلوب e کنترل می‌شود. لذا افراد با ژنوتیپ EE و Ee دارای لاله گوش نرمال و فرد دارای ژنوتیپ ee لاله گوش چسبیده خواهد داشت. فرض کنید ژنوتیپ هر کدام از والدین به صورت Ee است. در صورتی که این زوج دارای ٤ فرزند باشند، فرزندان با لاله گوش چسبیده را با x نشان داده و توزیع احتمال x را به دست آورید.


Ee×Ee
نر
Ee
E
e
ماده
E
EE
Ec
Ee
e
Ee
ee