فصل اول

مروری بر احتمال، متغیرهای تصادفی و توزیع‌های احتمال

۱.۱ مقدمه

بسیاری از مردم آمار را به عنوان ثبت و نمایش داده‌ها و نمودارهای مربوط به وضعیت‌های اقتصادی، اجتماعی یا سیاسی می‌شناسند. اما آمار به‌عنوان یک موضوع علمی، شامل مفاهیم و روش‌هایی برای جمع‌آوری داده‌ها به‌وسیلهٔ یک فرآیند آزمایش و مشاهده، تنظیم و خلاصه کردن آن‌ها و استنباط و نتیجه‌گیری به‌وسیلهٔ تحلیل این داده‌ها است. آمار استنباطی مجموعه‌ای از روش‌های علمی برای پیش‌بینی، برآورد، تصمیم‌گیری صحیح، به‌دست آوردن نتایج معتبر و ... است. این روش‌های علمی که به روش‌های آماری مشهور است در جهت نتیجه‌گیری، تعمیم و پیش‌بینی در مورد مشخصه‌های یک جامعه براساس نتایج به‌دست آمده از بخشی از جامعه (نمونه) بنا شده‌اند. نتایج به‌دست آمده از نمونه را نمی‌توان با اطمینان کامل به کل جامعه تعمیم داد، ولی می‌توان میزان صحت و دقت استنباط را با استفاده از قوانین احتمال مشخص کرد. بنابراین برای تعمیم نتایج نمونه به جامعه نیاز به مباحث احتمال، متغیرهای تصادفی و توزیع‌های احتمال داریم که در این فصل به‌طور مختصر این مفاهیم را مرور می‌کنیم. برای آشنایی بیشتر با این مفاهیم می‌توان به شلدون راس (۲۰۱۲)، والپول و همکاران (۲۰۱۲)، بهبودیان (۱۳۹۳)، پارسیان (۱۳۹۳) و نعمت‌الهی (۱۳۹۴) مراجعه نمود. در مورد جامعه، نمونه و روش‌های آماری به‌طور مفصل در فصل‌های بعد بحث خواهیم کرد.

۲.۱ احتمال

در این بخش مفاهیم احتمال را به‌طور مختصر مرور می‌کنیم.

تعریف ۱.۱ آزمایشی که در شرایط یکسان بتوان آن را تکرار کرد و نتیجهٔ آن قبل از انجام آزمایش تعیین شدنی نبوده ولی همهٔ نتایج ممکن آن قابل تعیین باشد، یک آزمایش تصادفی گویند.

تعریف ۲.۱ مجموعهٔ تمام نتایج یک آزمایش تصادفی را فضای نمونه گویند و آن را با نماد S نمایش می‌دهند.

مثال ۱.۱ در هر یک از موارد زیر فضای نمونهٔ آزمایش تصادفی مشخص شده است.

(الف) در پرتاب یک سکه اگر H نمایانگر مشاهدهٔ شیر و T نمایانگر مشاهدهٔ خط باشد

آنگاه فضای نمونه $S = \{H, T\}$ خواهد بود. در پرتاب دو سکه (یا دو مرتبه یک سکه) فضای نمونه به‌صورت $S = \{HH, HT, TH, TT\}$ است.

(ب) در پرتاب یک تاس فضای نمونه $S = \{۱, ۲, ۳, ۴, ۵, ۶\}$ است.

(پ) در آزمایش تصادفی مشاهدهٔ قطعه‌های تولیدی یک کارخانه تا رسیدن به یک قطعهٔ خراب، اگر D نمایانگر خراب بودن قطعه و N نمایانگر سالم بودن قطعه باشد آنگاه فضای نمونه $S = \{D, ND, NND, NNND, \cdots\}$ است. همچنین اگر تعداد قطعه‌های انتخابی تا مشاهدهٔ قطعه خراب مد نظر باشد، فضای نمونه $S = \{۱, ۲, ۳, ۴, \cdots\}$ خواهد بود.

(ت) در آزمایش تصادفی اندازه‌گیری طول دورهٔ درمان یک بیماری که حداقل یک و حداکثر پنج روز است، فضای نمونه $S = [۱, ۵]$ است.

با توجه به مثال ۱.۱، فضای نمونه می‌تواند گسسته (متناهی مانند حالت (الف) و (ب) و نامتناهی شمارش‌پذیر مانند حالت (پ)) یا پیوسته (یک فاصله از اعداد حقیقی مانند حالت (ت)، یا یک سطح در فضای دوبعدی و ...) باشد.

تعریف ۳.۱ در یک فضای نمونه گسسته هر زیر مجموعه از فضای نمونه را یک پیشامد گویند.

برای مثال در مثال ۱.۱.ب پیشامد عدد زوج عبارت از $A = \{۲, ۴, ۶\}$ است و در مثال ۱.۱.پ پیشامد اینکه قطعه تولیدی خراب حداکثر در ۳ مشاهده اول اتفاق افتد برابر $B = \{D, ND, NND\}$ یا $C = \{۱, ۲, ۳\}$ است.

گوییم پیشامد A در فضای نمونهٔ S رخ داده است، هرگاه نتیجهٔ آزمایش تصادفی منجر به مشاهدهٔ عضوی از پیشامد A گردد. اگر A و B دو پیشامد در فضای نمونه S باشند، وقوع همزمان آن‌ها عبارت از وقوع اشتراک آن‌ها یعنی $A \cap B = \{x | x \in A, x \in B\}$، وقوع حداقل یکی از این پیشامدها عبارت از وقوع اجتماع آن‌ها یعنی $A \cup B = \{x | x \in A, x \in B\}$ و رخ ندادن پیشامد A به‌معنای وقوع متمم پیشامد A یعنی $A' = \{x | x \notin A\}$ است.

تعریف ۴.۱ دو پیشامد A و B را ناسازگار (جدا) گویند هرگاه $A \cap B = \phi$، یعنی هر دو نتوانند همزمان اتفاق بیفتند. پیشامدهای $A_۱, A_۲, \cdots, A_n$ را دوبه‌دو ناسازگار گویند هر گاه برای هر $i \neq j$، $A_i \cap A_j = \phi$ باشد.

مثال ۲.۱ در مثال ۱.۱.ب اگر A پیشامد مشاهدهٔ عدد زوج، B پیشامد مشاهدهٔ عدد کمتر از ۴ و C پیشامد مشاهدهٔ عدد فرد در پرتاب تاس باشند، آنگاه $A \cap B = \{۲\}$، $A \cup B = \{۱, ۲, ۳, ۴, ۶\}$ و $A \cap C = \phi$، یعنی A و C ناسازگار هستند.

۱.۲.۱ احتمال

تعریف ۵.۱ احتمال تابعی از مجموعهٔ متشکل از زیرمجموعه‌های فضای نمونه به داخل مجموعهٔ اعداد حقیقی است. به‌عبارت دقیق‌تر احتمال تابعی مانند P است که به هر پیشامد A از فضای نمونه S عدد حقیقی $P(A)$ را به‌گونه‌ای نسبت می‌دهد که در سه اصل موضوع زیر صدق کند.

۱) $P(S) = ۱$،

۲) برای هر پیشامد A در S، $P(A) \geq ۰$،

۳) اگر $A_۱, A_۲, A_۳, ...$ پیشامدهای دو به دو ناسازگار باشند، آنگاه

$$P(A_۱ \cup A_۲ \cup A_۳ \cup ...) = P(A_۱) + P(A_۲) + P(A_۳) + ...$$

اگر فضای نمونه S یک فضای گسسته به‌صورت $S = \{e_۱, e_۲, e_۳, ...\}$ باشد، مدل احتمال روی این فضا عبارت است از نسبت دادن وزن‌های (احتمال‌های) نامنفی $p_۱, p_۲, ...$ به نقاط فضای نمونهٔ S به‌طوری که مجموع تمام این اعداد برابر یک شود، یعنی

S	$e_۱$	$e_۲$	$e_۳$	...
احتمال	$p_۱$	$p_۲$	$p_۳$	...

$$۰ \leq p_i \leq ۱, i = ۱, ۲, ..., n, \quad \sum_{i=۱}^{\infty} p_i = ۱$$

احتمال وقوع یک پیشامد A در این فضا برابر مجموع وزن‌های نسبت داده شده به نقاط تشکیل دهنده A است، یعنی

$$A = \{e_۱, e_۲, e_۳, ..., e_m\} \implies P(A) = p_۱ + p_۲ + ... + p_m = \sum_{i=۱}^{m} p_i$$

مثال ۳.۱ در پرتاب یک تاس فرض کنید A پیشامد مشاهدهٔ عدد کمتر از ۴ باشد. دراین‌صورت

(الف) اگر هیچ‌گونه اطلاعی از شانس وقوع اعداد روی تاس نداشته باشیم آنگاه طبیعی است که شانس وقوع هر عدد را یکسان در نظر بگیریم. پس مدل احتمال برابر است

با

S	۱	۲	۳	۴	۵	۶
احتمال	$\frac{۱}{۶}$	$\frac{۱}{۶}$	$\frac{۱}{۶}$	$\frac{۱}{۶}$	$\frac{۱}{۶}$	$\frac{۱}{۶}$

حال اگر $A = \{۱, ۲, ۳\}$ آنگاه $P(A) = \frac{۱}{۶} + \frac{۱}{۶} + \frac{۱}{۶} = \frac{۳}{۶} = \frac{۱}{۲}$.

(ب) اگر در این تاس شانس مشاهدهٔ زوج دو برابر فرد باشد آنگاه شانس هر عدد فرد w و شانس هر عدد زوج $۲w$ خواهد بود. چون جمع احتمال‌ها برابر یک است پس بایستی $۹w = ۱$ یا $w = \frac{۱}{۹}$ پس مدل احتمال برابر است با

S	۱	۲	۳	۴	۵	۶
احتمال	$\frac{۱}{۹}$	$\frac{۲}{۹}$	$\frac{۱}{۹}$	$\frac{۲}{۹}$	$\frac{۱}{۹}$	$\frac{۲}{۹}$

و در نتیجه $P(A) = \frac{۱}{۹} + \frac{۲}{۹} + \frac{۱}{۹} = \frac{۴}{۹}$.

مدل احتمال یکنواخت در مثال ۳.۱.الف مشاهده شد موقعی که نقاط فضای نمونه شانس مساوی برای وقوع داشته باشند آنگاه

$$P(A) = \frac{n(A)}{n(S)} = \frac{\text{تعداد اعضای پیشامد } A}{\text{تعداد اعضای فضای نمونه } S} = \frac{\text{تعداد حالت‌های مطلوب}}{\text{تعداد کل حالت‌ها}} \qquad (۱.۱)$$

این مدل را مدل احتمال یکنواخت گویند. در این حالت نیاز به شمارش تعداد اعضای یک مجموعه داریم که در زیر به‌طور خلاصه قواعد شمارش اعضای یک مجموعه را مرور می‌کنیم.

قواعد شمارش ترتیبی که در آن می‌توان اعضای یک مجموعه را کنار یکدیگر قرار داد یک جایگشت گویند.

(الف) اگر n عنصر متمایز را بخواهیم در یک صف کنار یکدیگر قرار دهیم آنگاه تعداد جایگشت‌های ممکن برابر است با

$$n! = n(n-۱)(n-۲)\cdots(۲)(۱) = n!$$

(ب) اگر n عنصر وجود داشته باشد که $n_۱$ تای آن‌ها از نوع اول، $n_۲$ تای آن‌ها از نوع دوم، ... و n_r تای آن‌ها از نوع r ام باشند که $n_۱ + n_۲ + \ldots + n_r = n$، آنگاه تعداد جایگشت‌های این عناصر برابر $\frac{n!}{n_۱!n_۲!\ldots n_r!} = \binom{n_۱, n_۲, \ldots, n_r}{n}$ است.

(پ) اگر از بین n عنصر متمایز بخواهیم r عنصر را انتخاب کرده و در یک صف قرار دهیم، دراین‌صورت

(پ.۱) اگر تکرار عناصر مجاز نباشد و ترتیب قرار گرفتن آن‌ها در کنار یکدیگر مهم باشد، آنگاه تعداد راه‌های ممکن برابر $P_r^n = \frac{n!}{(n-r)!}$ است که به آن تبدیل (جایگشت) r از n گویند.

(پ.۲) اگر تکرار عناصر مجاز باشد و ترتیب قرار گرفتن آن‌ها در کنار یکدیگر مهم باشد، آنگاه تعداد راه‌های ممکن برابر n^r است.

(پ.۳) اگر تکرار عناصر مجاز نباشد و ترتیب قرار گرفتن آن‌ها در کنار یکدیگر مهم نباشد، آنگاه تعداد راه‌های ممکن برابر $\frac{n!}{r!(n-r)!} = \binom{n}{r}$ است که به آن ترکیب r از n گویند.

(پ.۴) اگر تکرار عناصر مجاز باشد و ترتیب قرار گرفتن آن‌ها در کنار یکدیگر مهم نباشد، آنگاه تعداد راه‌های ممکن برابر $\binom{n+r-1}{r}$ است.

مثال ۴.۱ فرض کنید در یک کتاب‌فروشی ۳ کتاب ریاضی عمومی، ۵ کتاب فیزیک و ۴ کتاب آمار وجود دارد.

(الف) اگر بخواهیم کتاب‌ها را در کنار یکدیگر در قفسه قرار دهیم احتمال اینکه هر ۳ کتاب ریاضی پهلوی هم قرار گیرند را بیابید.

(ب) اگر ۶ کتاب از بین این کتاب‌ها انتخاب کنیم احتمال اینکه شامل ۲ کتاب ریاضی و ۳ کتاب فیزیک باشد را بیابید. احتمال اینکه حداکثر یک کتاب ریاضی انتخاب شود را بیابید.

حل : در این مثال ۳ کتاب ریاضی عمومی مانند ۳ حرف M، ۵ کتاب فیزیک مانند ۵ حرف P و ۴ کتاب آمار مانند ۴ حرف S هستند.

(الف) تعداد کل راه‌های قرار گرفتن کتاب‌ها در کنار یکدیگر $n(S) = \frac{۱۲!}{۳!۵!۴!}$ است. حال ۳ کتاب ریاضی MMM به‌صورت یک گروه متصل در کنار یکدیگر قرار می‌گیرند و اگر A پیشامد قرار گرفتن آن‌ها در کنار یکدیگر باشد آنگاه $n(A) = \frac{۱۰!}{۱!۵!۴!}$ و در نتیجه

$$P(A) = \frac{n(A)}{n(S)} = \frac{۱۰!}{۱!۵!۴!} \times \frac{۳!۵!۴!}{۱۲!} = \frac{۳!۱۰!}{۱۲!} = \frac{۱}{۲۲}$$

(ب) تعداد کل راه‌های انتخاب ۶ کتاب $n(S) = \binom{۱۲}{۶}$ است. اگر B پیشامد انتخاب ۲ کتاب ریاضی و ۳ فیزیک باشد آنگاه $\dfrac{\binom{۳}{۲}\binom{۵}{۳}\binom{۴}{۱}}{\binom{۱۲}{۶}} = \dfrac{n(B)}{n(S)} = P(A)$ که $\binom{۴}{۱}$ مربوط به انتخاب یک کتاب باقی‌مانده از بین کتاب‌های آمار است. اگر C پیشامد انتخاب حداکثر یک کتاب ریاضی باشد آنگاه

$$P(C) = \frac{\binom{۳}{۰}\binom{۹}{۶}}{\binom{۱۲}{۶}} + \frac{\binom{۳}{۱}\binom{۹}{۵}}{\binom{۱۲}{۶}}$$

چند قانون احتمال با استفاده از سه اصل احتمال می‌توان نتایج زیر را به‌دست آورد.

(الف) $P(\phi) = ۰$

(ب) اگر A و B دو پیشامد دلخواه باشند آنگاه

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A - B) = P(A) - P(A \cap B)$$

و اگر A و B دو پیشامد ناسازگار باشند آنگاه $P(A \cup B) = P(A) + P(B)$

(پ) $P(A') = ۱ - P(A)$

احتمال در فضای نمونهٔ پیوسته در یک حالت خاص فضای نمونه پیوسته به‌صورت یک فاصله کراندار $S = [a, b]$ از اعداد حقیقی (یا یک سطح در فضای دو بعدی یا ...) است و پیشامدها به‌صورت یک زیرفاصله یا اجتماعی از زیر فاصله‌ها (یا یک زیر سطح در فضای دو بعدی یا ...) هستند. در این حالت احتمال هر پیشامد A را می‌توان به‌صورت زیر محاسبه کرد.

$$P(A) = \frac{\text{طول زیر فاصله } A}{\text{طول فاصله } S} \quad \left(\text{یا } = \frac{\text{مساحت ناحیه } A}{\text{مساحت ناحیه } S} \ \text{یا...}\right) \qquad (۲.۱)$$

مثال ۵.۱ عددی را به تصادف از فاصلهٔ اعداد حقیقی $[۲, ۵]$ انتخاب می‌کنیم.

(الف) احتمال اینکه عدد انتخابی در فاصله $[۳, ۴/۵]$ باشد را بیابید.

(ب) احتمال اینکه عدد انتخابی دقیقاً ۳ باشد را بیابید.

حل :

(الف) در این حالت $S = [۲, ۵]$ و $A = [۳, ۴/۵]$، پس از (۲.۱) داریم

$$P(A) = \frac{A \text{ طول فاصله}}{S \text{ طول فاصله}} = \frac{۴/۵ - ۳}{۵ - ۲} = \frac{۱/۵}{۳} = \frac{۱}{۲}$$

(ب) چون فاصله $[۲, ۵]$ شامل بی‌نهایت عدد است و انتخاب یک عدد خاص در این فاصله غیر ممکن است، پس $P(\{۳\}) = ۰$.

۲.۲.۱ احتمال شرطی

در برخی مسئله‌ها می‌خواهیم بعد از وقوع یک پیشامد A، احتمال وقوع پیشامد B را به‌دست آوریم. این احتمال را با $P(B|A)$ نمایش می‌دهند و به‌صورت زیر تعریف می‌شود.

تعریف ۶.۱ احتمال وقوع پیشامد B به شرط وقوع پیشامد A به‌صورت زیر تعریف می‌شود.

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \qquad P(A) \neq ۰ \tag{۳.۱}$$

مثال ۶.۱ دانشجویان سال اول دانشکده‌ای، ۲۵% از درس ریاضی، ۱۵% از درس فیزیک و ۱۰% از هر دو درس ریاضی و فیزیک مردود شده‌اند. اگر یک دانشجو از درس ریاضی مردود شده باشد، احتمال اینکه از درس فیزیک نیز مردود شده باشد را بیابید.

حل : اگر M و P به‌ترتیب پیشامدهای مردود شدن از درس‌های ریاضی و فیزیک باشد، آنگاه $P(M) = ۰/۲۵$، $P(P) = ۰/۱۵$ و $P(M \cap P) = ۰/۱۰$. بنابراین $P(P|M) = \frac{P(P \cap M)}{P(M)} = \frac{۰/۱۰}{۰/۲۵} = ۰/۴$.

قانون ضرب احتمال اگر A و B دو پیشامد باشند که بتوانند همزمان اتفاق بیفتند، آنگاه از رابطهٔ (۳.۱) نتیجه می‌شود که

$$P(A \cap B) = P(A)P(B|A) \tag{۴.۱}$$

این فرمول را قانون ضرب احتمال گویند. در حالت کلی اگر $A_۱, A_۲, ..., A_n$ پیشامدهایی باشند که بتوانند همزمان اتفاق بیفتند، آنگاه

$$\begin{aligned} P(A_۱ \cap A_۲ \cap A_۳ \cap ... \cap A_n) &= P(A_۱)P(A_۲|A_۱)P(A_۳|A_۱ \cap A_۲) \\ &\quad ...P(A_n|A_۱ \cap A_۲ \cap ... \cap A_{n-۱}) \end{aligned} \tag{۵.۱}$$

مثال ۷.۱ در ظرفی ۳ مهره سفید و ۵ مهره سیاه وجود دارد. از این ظرف سه مهره یک به یک و بدون جای‌گذاری انتخاب می‌کنیم.

(الف) احتمال اینکه مهرهٔ انتخابی اول سفید و مهرهٔ انتخابی دوم سیاه باشد را بیابید.

(ب) احتمال اینکه مهره‌های انتخابی اول و سوم سیاه و مهرهٔ انتخابی دوم سفید باشد را بیابید.

حل : اگر W_i و B_i به‌ترتیب پیشامد انتخاب مهره سفید و مهره سیاه در انتخاب i ام، $i = ۱, ۲, ۳$ باشند، آنگاه

(الف)
$$P(W_۱ \cap B_۲) = P(W_۱)P(B_۲|W_۱) = \frac{۳}{۸} \times \frac{۵}{۷} = \frac{۱۵}{۵۶}$$

(ب)
$$P(B_۱ \cap W_۲ \cap B_۳) = P(B_۱)P(W_۲|B_۱)P(B_۳|B_۱ \cap W_۲) = \frac{۵}{۸} \times \frac{۳}{۷} \times \frac{۴}{۶} = \frac{۱۰}{۵۶}$$

پیشامدهای مستقل در مثال ۷.۱ فرض کنید انتخاب مهره‌ها با جای‌گذاری باشد. دراین‌صورت انتخاب مهرهٔ اول تاثیری در انتخاب مهرهٔ دوم ندارد و $P(B_۲|W_۱) = \frac{۵}{۸} = P(B_۲)$. در این حالت پیشامدهای $W_۱$ و $B_۲$ را مستقل گویند.

تعریف ۷.۱ دو پیشامد A و B را مستقل گویند اگر و فقط اگر

$$P(A \cap B) = P(A)P(B) \qquad (۶.۱)$$

همچنین سه پیشامد A، B و C را مستقل گویند اگر و فقط اگر رابطه‌های زیر برقرار باشند.

$$\begin{aligned} &P(A \cap B) = P(A)P(B)\ , && P(A \cap C) = P(A)P(C) \\ &P(B \cap C) = P(B)P(C)\ , && P(A \cap B \cap C) = P(A)P(B)P(C) \end{aligned} \qquad (۷.۱)$$

برای مثال، در مثال ۶.۱ اگر انتخاب مهره‌ها با جای‌گذاری باشد آنگاه

$$P(W_۱ \cap B_۲) = P(W_۱)P(B_۲) = \frac{۳}{۸} \times \frac{۵}{۸} = \frac{۱۵}{۶۴}$$
$$P(B_۱ \cap W_۲ \cap B_۳) = P(B_۱)P(W_۲)P(B_۳) = \frac{۵}{۸} \times \frac{۳}{۸} \times \frac{۵}{۸} = \frac{۷۵}{۵۱۲}.$$

۳.۱ متغیرهای تصادفی

در یک آزمایش تصادفی اغلب به‌جای نتیجهٔ به‌دست آمده از آزمایش تصادفی به مقدار عددی که به این نتیجه نسبت داده شده است علاقه‌مند هستیم. این مقادیر عددی که از تعریف یک تابع روی فضای نمونه به‌دست می‌آیند متغیر تصادفی نامیده می‌شوند.

تعریف ۸.۱ یک متغیر تصادفی تابعی از فضای نمونه به مجموعهٔ اعداد حقیقی است به‌طوری که به هر عضو از فضای نمونه یک عدد حقیقی را نسبت می‌دهد.

برای نمایش متغیر تصادفی یکی از حروف بزرگ مانند X، Y و ... و برای نمایش مقادیری که متغیر تصادفی می‌تواند اختیار کند از حروف کوچک معادل آن یعنی x، y و ... استفاده می‌شود. مجموعهٔ مقادیر و یا برد متغیر تصادفی را با S_X نمایش می‌دهند و آن را تکیه‌گاه X می‌نامند.

احتمال هر پیشامد را می‌توان به‌وسیلهٔ احتمال وقوع برخی مقادیر متغیر تصادفی نمایش داد. برای مثال احتمال $(X \leq ۲)$ به معنای احتمال پیشامد $\{w \in S | X(w) \leq ۲\}$ است و آن را با نماد $P(X \leq ۲)$ نمایش می‌دهند.

مثال ۸.۱ فرض کنید سکه‌ای را سه مرتبه پرتاب کنیم و متغیر تصادفی X را برابر تعداد شیرهای مشاهده شده در سه پرتاب سکه در نظر بگیریم. احتمال اینکه حداقل ۲ شیر مشاهده کنیم را بیابید.

حل : فهرست نقاط فضای نمونه، مقادیری که متغیر تصادفی X به این نقاط نسبت می‌دهد و احتمال‌های مربوطه عبارت است از

(۸.۱)

S	TTT	TTH	THT	HTT	THH	HTH	HHT	HHH
x	۰	۱	۱	۱	۲	۲	۲	۳
احتمالات	$\frac{۱}{۸}$	$\frac{۱}{۸}$	$\frac{۱}{۸}$	$\frac{۱}{۸}$	$\frac{۱}{۸}$	$\frac{۱}{۸}$	$\frac{۱}{۸}$	$\frac{۱}{۸}$

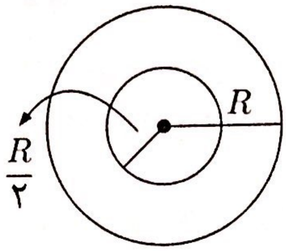
از جدول ۸.۱ احتمال مشاهدهٔ حداقل ۲ شیر برابر است با

$$P(X \geq ۲) = P(X = ۲) + P(X = ۳) = \frac{۳}{۸} + \frac{۱}{۸} = \frac{۱}{۲}$$

مثال ۹.۱ از داخل دایره‌ای به شعاع R نقطه‌ای به تصادف انتخاب می‌کنیم و متغیر تصادفی Y را برابر فاصلهٔ نقطهٔ انتخابی تا مرکز دایره در نظر می‌گیریم. احتمال اینکه نقطه به مرکز نزدیک‌تر باشد تا به محیط دایره را بیابید.

حل : در اینجا $S = [۰, R]$ و $P\left(Y < \frac{R}{۲}\right)$ مورد سوال است. بنابراین از روی شکل ۱.۱ داریم

$$P\left(Y < \frac{R}{۲}\right) = \frac{\text{مساحت دایره به شعاع } \frac{R}{۲}}{\text{مساحت دایره به شعاع } R} = \frac{\pi\left(\frac{R}{۲}\right)^{۲}}{\pi R^{۲}} = \frac{۱}{۴}$$



شکل ۱.۱

متغیرهای تصادفی گسسته و پیوسته متغیر تصادفی که مجموعهٔ مقادیر آن متناهی یا نامتناهی شمارش‌پذیر باشد (مثال ۸.۱) را متغیر تصادفی گسسته و متغیر تصادفی که مجموعهٔ مقادیر آن یک فاصله یا اجتماع چند فاصله عددی باشد (مثال ۹.۱) را متغیر تصادفی پیوسته گویند.

۱.۳.۱ توزیع احتمالات گسسته

در مثال ۸.۱ متغیر تصادفی X دارای تکیه‌گاه $S_X = \{۰, ۱, ۲, ۳\}$ است و از (۸.۱) می‌توان احتمال‌های زیر را به‌دست آورد.

x	۰	۱	۲	۳
$P(X = x)$	$\frac{۱}{۸}$	$\frac{۳}{۸}$	$\frac{۳}{۸}$	$\frac{۱}{۸}$

تابع $P(X = x)$ را تابع احتمال متغیر تصادفی گسسته X گویند و خواص آن در تعریف زیر آورده شده است.

تعریف ۹.۱ تابع $f_X(x) = P(X = x)$ را تابع احتمال متغیر تصادفی گسسته X گویند هرگاه

(الف) برای هر $x \in R$، $f_X(x) \geq ۰$،

(ب) $\sum_{x \in R} f_X(x) = ۱$.

برای محاسبهٔ احتمال پیشامد $(X \in C)$ که C زیر مجموعه‌ای از اعداد حقیقی است، از رابطهٔ زیر استفاده می‌کنیم.

$$P(X \in C) = \sum_{x \in C} f_X(x) \tag{۹.۱}$$

مثال ۱۰.۱ سکه‌ای را که شانس مشاهدهٔ شیر در آن دو برابر خط است آن‌قدر پرتاب می‌کنیم تا یک خط مشاهده شود. اگر متغیر تصادفی X برابر تعداد پرتاب‌های لازم تا رسیدن به یک خط باشد، تابع احتمال X و احتمال اینکه حداقل ۴ پرتاب لازم باشد را بیابید.

حل : فهرست نقاط فضای نمونه و تابع احتمال X به‌صورت زیر است.

S	T	HT	HHT	$HHHT$	...
x	۱	۲	۳	۴	...
$f_X(x) = P(X = x)$	$\frac{۱}{۳}$	$\left(\frac{۲}{۳}\right)\left(\frac{۱}{۳}\right)$	$\left(\frac{۲}{۳}\right)^{۲}\left(\frac{۱}{۳}\right)$	$\left(\frac{۲}{۳}\right)^{۳}\left(\frac{۱}{۳}\right)$	...

$$f_X(x) = \left(\frac{۲}{۳}\right)^{x-۱} \frac{۱}{۳}, \quad x = ۱, ۲, \ldots$$

همان‌گونه که ملاحظه می‌شود با استفاده از مجموع یک سری هندسی داریم

$$\sum_{x=۱}^{\infty} f_X(x) = \sum_{x=۱}^{\infty} \left(\frac{۲}{۳}\right)^{x-۱} \frac{۱}{۳} = \frac{۱}{۳} \times \frac{۱}{۱ - \frac{۲}{۳}} = ۱$$

$$P(X \geq ۴) = \sum_{x=۴}^{\infty} \left(\frac{۲}{۳}\right)^{x-۱} \frac{۱}{۳} = \frac{۱}{۳} \times \frac{\left(\frac{۲}{۳}\right)^{۳}}{۱ - \frac{۲}{۳}} = \frac{۸}{۲۷}$$

تابع توزیع (تجمعی)

تعریف ۱۰.۱ اگر X یک متغیر تصادفی گسسته با تابع احتمال $f_X(x)$ باشد آنگاه تابع توزیع(تجمعی) X برای هر $x \in R$ به‌صورت زیر تعریف می‌شود

$$F_X(x) = P(X \leq x) = \sum_{t \leq x} f_X(t) \tag{۱۰.۱}$$

مثال ۱۱.۱ سه مهره به شماره‌های ۱، ۲ و ۳ را در ۳ جعبه به شماره‌های ۱، ۲ و ۳ به‌طور تصادفی می‌ریزیم به‌گونه‌ای که هر جعبه شامل تنها یک مهره باشد. اگر متغیر تصادفی X

برابر تعداد جورها (تعداد مهره‌هایی که در جعبه با شماره متناظر خودشان قرار گرفته‌اند) باشد، تابع احتمال و تابع توزیع X را به‌دست آورید.

حل : در این مثال $S_X = \{۰, ۱, ۳\}$ و به $۳! = ۶$ طریق ۳ مهره در ۳ جعبه قرار می‌گیرند. بنابراین

$$P(X = ۱) = P(\{۱۳۲, ۳۲۱, ۲۱۳\}) = \frac{۳}{۶}$$

به همین ترتیب تابع احتمال X به‌صورت زیر به‌دست می‌آید

x	۰	۱	۳
$f_X(x) = P(X = x)$	$\frac{۲}{۶}$	$\frac{۳}{۶}$	$\frac{۱}{۶}$

در نتیجه تابع توزیع X به‌صورت زیر به‌دست می‌آید.

$$F_X(x) = P(X \leq x) = \begin{cases} ۰ & x < ۰ \\ \frac{۲}{۶} & ۰ \leq x < ۱ \\ \frac{۲}{۶} + \frac{۳}{۶} = \frac{۵}{۶} & ۱ \leq x < ۳ \\ ۱ & ۳ \leq x \end{cases}$$

خواص تابع توزیع با توجه به تعریف تابع توزیع و مثال ۱۱.۱، خواص زیر را برای تابع توزیع می‌توان بیان کرد

(الف) برای هر $x \in R$، $۰ \leq F_X(x) \leq ۱$.

(ب) تابع توزیع یک تابع غیرنزولی است، یعنی برای هر $x_۱ < x_۲$،

$$F_X(x_۱) \leq F_X(x_۲)$$

(پ)

$$F_X(-\infty) = \lim_{x \to -\infty} F_X(x) = ۰ \quad , \quad F_X(+\infty) = \lim_{x \to +\infty} F_X(x) = ۱$$

(ت) تابع توزیع همواره در هر نقطه از سمت راست پیوسته است، یعنی

$$\lim_{x \to a^+} F_X(x) = F_X(a)$$

توجه کنید که از روی تابع توزیع یک متغیر تصادفی گسسته می‌توان تابع احتمال و همچنین احتمال‌ها را به‌صورت زیر محاسبه کرد

$$f_X(x) = F_X(x) - F_X(x^-) = F_X(x) - \lim_{t \to x^-} F_X(t) \qquad (۱۱.۱)$$
$$P(a < x \leq b) = F_X(b) - F_X(a)$$

۲.۳.۱ توزیع احتمالات پیوسته

اگر X یک متغیر تصادفی پیوسته باشدآنگاه برای هر عدد حقیقی $b \in R$، $P(X = b) = ۰$ خواهد بود. بنابراین در این حالت تابع احتمال به‌صورت تعریف ۹.۱ وجود ندارد. برای متغیر تصادفی پیوسته به‌جای تابع احتمال، تابع چگالی احتمال وجود دارد که می‌توان آن را از روی تابع توزیع به‌صورت زیر به‌دست آورد.

همانند ۱۰.۱، تابع توزیع متغیر تصادفی پیوسته X به‌صورت زیر تعریف می‌شود

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)dt \tag{۱۲.۱}$$

تابع $f_X(x)$ که در رابطهٔ (۱۲.۱) صدق می‌کند را تابع چگالی احتمال و تابع $F_X(x)$ در این رابطه را تابع توزیع متغیر تصادفی پیوسته X گویند. تابع توزیع متغیر تصادفی پیوسته همان خواص تابع توزیع متغیر تصادفی گسسته را دارد و علاوه بر آن در هر نقطه این تابع پیوسته است. از رابطه (۱۲.۱) نتیجه می‌شود که

$$f_X(x) = \frac{d}{dx}F_X(x) = F'_X(x) \tag{۱۳.۱}$$

خواص تابع چگالی احتمال تابع چگالی احتمال یک متغیر تصادفی پیوسته X خواص زیر را دارد

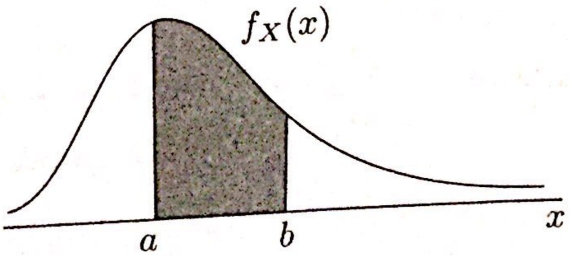
(الف) برای هر $x \in R$، $f_X(x) \geq ۰$،

(ب) $\int_{-\infty}^{+\infty} f_X(x)dx = ۱$

و برای محاسبه احتمال‌ها با استفاده از تابع چگالی احتمال و تابع توزیع، رابطهٔ زیر را داریم.

$$P(a < x \leq b) = \int_{a}^{b} f_X(x)dx = F_X(b) - F_X(a) \tag{۱۴.۱}$$

احتمال بالا سطح زیر منحنی $f_X(x)$ بین خطوط $x = a$ و $x = b$ است.



شکل ۲.۱

مثال ۱۲.۱ تابع چگالی احتمال متغیر تصادفی X به‌صورت زیر داده شده است.

$$f_X(x) = \begin{cases} kx^{۲} & -۱ < x < ۱ \\ ۰ & \text{سایر نقاط} \end{cases}$$

(الف) مقدار k را تعیین کنید.

(ب) تابع توزیع X را به‌دست آورید و $P(|X| \leq \frac{۱}{۲})$ و $P([X] = -۱)$ را محاسبه کنید.

حل : (الف) برای تعیین k توجه کنید که

$$۱ = \int_{-\infty}^{+\infty} f_X(x)dx = \int_{-۱}^{۱} kx^{۲}dx = \frac{k}{۳}x^{۳}\Big|_{-۱}^{۱} = \frac{۲k}{۳}$$

پس $k = \frac{۳}{۲}$.

(ب) با توجه به (۱۲.۱) داریم

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \begin{cases} ۰ & x < -۱ \\ \int_{-۱}^{x} \frac{۳}{۲}t^{۲}dt & -۱ \leq x < ۱ \\ ۱ & ۱ \leq x \end{cases}$$

$$= \begin{cases} ۰ & x < -۱ \\ \frac{۱}{۲}(x^{۳} + ۱) & -۱ \leq x < ۱ \\ ۱ & ۱ \leq x \end{cases}$$

همچنین

$$\begin{aligned} P\left(|X| \leq \frac{۱}{۲}\right) &= P\left(-\frac{۱}{۲} \leq X \leq \frac{۱}{۲}\right) = F_X\left(\frac{۱}{۲}\right) - F_X\left(-\frac{۱}{۲}\right) \\ &= \frac{۱}{۲}\left(\frac{۹}{۸} - \frac{۷}{۸}\right) = \frac{۱}{۸} \\ P([X] = -۱) &= P(-۱ \leq X < ۰) = \int_{-۱}^{۰} f_X(x)dx = \int_{-۱}^{۰} \frac{۳}{۲}x^{۲}dx = \frac{۱}{۲} \end{aligned}$$

۳.۳.۱ توزیع احتمالات توأم دو متغیره

در برخی از آزمایش‌های تصادفی ممکن است نتیجهٔ دو یا چند متغیر تصادفی به‌طور همزمان مد نظر باشد. در این حالت توزیع احتمالات دو یا چند متغیر تصادفی، به‌صورت تعمیمی از

حالت یک متغیر تصادفی به‌دست می‌آید. برای مثال در حالت گسسته توزیع احتمال توأم برای وقوع همزمان دو متغیر تصادفی X و Y به‌صورت زیر تعریف می‌شود

$$f_{X,Y}(x,y) = P(X = x, Y = y) \tag{۱۵.۱}$$

و به معنای وقوع همزمان دو پیشامد $(X = x)$ و $(Y = y)$ است. این تابع دارای خواص زیر است.

تعریف ۱۱.۱ تابع $f_{X,Y}(x,y)$ را یک تابع احتمال متغیرهای تصادفی گسسته X و Y گویند هرگاه

(الف) برای هر x و y، $f_{X,Y}(x,y) \geq ۰$،

(ب) $\sum\limits_{x}\sum\limits_{y} f_X(x,y) = ۱$.

برای محاسبهٔ احتمال قرار گرفتن (X,Y) در ناحیهٔ A در صفحهٔ xy از رابطهٔ زیر استفاده می‌کنیم

$$P((X,Y) \in A) = \sum\sum_{(x,y)\in A} f_{X,Y}(x,y) \tag{۱۶.۱}$$

مثال ۱۳.۱ از یک ظرف میوه که شامل ۳ پرتقال، ۲ سیب و ۳ موز است چهار میوه را به تصادف انتخاب می‌کنیم. اگر X تعداد پرتقال‌های انتخابی و Y تعداد سیب‌های انتخابی باشند، تابع احتمال توأم X و Y را به‌دست آورید و $P(۲X + Y < ۳)$ را محاسبه کنید.

حل : در این مثال $S_X = \{۰, ۱, ۲, ۳\}$ و $S_Y = \{۰, ۱, ۲\}$ است. همچنین

$$f_{X,Y}(۱,۱) = P(X = ۱, Y = ۱) = \frac{\binom{۳}{۱}\binom{۲}{۱}\binom{۳}{۲}}{\binom{۸}{۴}} = \frac{۱۸}{۷۰}$$

$$f_{X,Y}(۳,۰) = P(X = ۳, Y = ۰) = \frac{\binom{۳}{۳}\binom{۲}{۰}\binom{۳}{۱}}{\binom{۸}{۴}} = \frac{۳}{۷۰}$$

و با انجام محاسبات مربوط به نقاط دیگر، جدول توزیع احتمالات توأم (X,Y) را به‌صورت زیر به‌دست می‌آوریم.

$$
\begin{array}{c|cccc|}
y \backslash x & ۰ & ۱ & ۲ & ۳ \\
\hline
۰ & ۰ & \frac{۳}{۷۰} & \frac{۹}{۷۰} & \frac{۳}{۷۰} \\
۱ & \frac{۲}{۷۰} & \frac{۱۸}{۷۰} & \frac{۱۸}{۷۰} & \frac{۲}{۷۰} \\
۲ & \frac{۳}{۷۰} & \frac{۹}{۷۰} & \frac{۳}{۷۰} & ۰ \\
\hline
\end{array}
\qquad (۱۷.۱)
$$

با توجه به جدول (۱۷.۱) داریم که

$$
\begin{aligned}
P(۲X+Y<۳) &= f_{X,Y}(۰,۰)+f_{X,Y}(۰,۱)+f_{X,Y}(۰,۲)+f_{X,Y}(۱,۰) \\
&= ۰+\frac{۲}{۷۰}+\frac{۳}{۷۰}+\frac{۳}{۷۰}=\frac{۸}{۷۰}
\end{aligned}
$$

برای متغیرهای تصادفی پیوسته نیز تعریف زیر را داریم.

تعریف ۱۲.۱ تابع $f_{X,Y}(x,y)$ را یک تابع چگالی احتمال توأم متغیرهای تصادفی پیوسته X و Y گویند هرگاه

(الف) برای هر x و y، $f_{X,Y}(x,y)\geq ۰$،

(ب) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{X,Y}(x,y)dxdy=۱$

برای محاسبهٔ احتمال قرار گرفتن (X,Y) در یک ناحیهٔ A در صفحه xy از رابطهٔ زیر استفاده می‌کنیم.

$$
P((X,Y)\in A)=\int_A\int f_{X,Y}(x,y)dxdy \qquad (۱۸.۱)
$$

مثال ۱۴.۱ تابع چگالی احتمال توأم متغیرهای تصادفی X و Y به‌صورت زیر است.

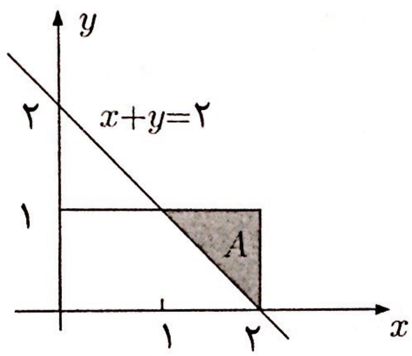
$$
f_{X,Y}(x,y)=\begin{cases} cx(۱+۳y^۲) & ۰<x<۲, ۰<y<۱ \\ ۰ & \text{سایر نقاط} \end{cases}
$$

مقدار c را تعیین کنید و $P(X+Y\geq ۲)$ را محاسبه کنید.

حل : برای محاسبه c از تعریف ۱۲.۱ داریم

$$\begin{aligned}
1 &= \int_0^1 \int_0^2 cx(1+3y^2)dxdy = \int_0^1 c(1+3y^2)\left[\frac{1}{2}x^2\right]_0^2 dy \\
&= 2c\int_0^1 (1+3y^2)dy = 2c\Big[y+y^3\Big]_0^1 = 4c
\end{aligned}$$

بنابراین $c=\frac{1}{4}$. با توجه به نمودار زیر داریم



شکل ۳.۱

$$\begin{aligned}
A &= \{(x,y)|0<y<1,\;\; 2-y<x<2\} \\
P(X+Y\geq 2) &= \int_A\int f_{X,Y}(x,y)dxdy = \int_0^1\int_{2-y}^2 \frac{1}{4}x(1+3y^2)dxdy \\
&= \int_0^1 \frac{1}{4}(1+3y^2)\left(2y-\frac{1}{2}y^2\right)dy \\
&= \frac{1}{4}\left[y^2+\frac{3}{2}y^4-\frac{1}{6}y^3-\frac{3}{10}y^5\right]_0^1 = \frac{61}{120}
\end{aligned}$$

۴.۳.۱ توزیع احتمالات حاشیه‌ای و شرطی

با داشتن تابع احتمال (چگالی احتمال) توأم متغیرهای تصادفی X و Y می‌توان تابع احتمال (چگالی احتمال) X به تنهایی و Y به تنهایی را محاسبه کرد که به آن‌ها تابع‌های احتمال (چگالی احتمال) حاشیه‌ای گویند. اگر X و Y دو متغیر تصادفی گسسته باشند، تابع احتمال حاشیه‌ای X و تابع احتمال حاشیه‌ای Y به‌صورت زیر به‌دست می‌آیند

$$f_X(x)=\sum_y f_{X,Y}(x,y), \qquad f_Y(y)=\sum_x f_{X,Y}(x,y) \qquad (19.1)$$

به‌همین ترتیب تابع چگالی احتمال حاشیه‌ای دو متغیر تصادفی پیوسته X و Y به‌صورت زیر به‌دست می‌آیند.

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy, \qquad f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dx \tag{۲۰.۱}$$

با استفاده از فرمول احتمال شرطی (۳.۱)، یعنی $P(A) \neq ۰$,$P(B|A) = \frac{P(A\cap B)}{P(A)}$، اگر قرار دهیم $B \equiv (X = x)$ و $A \equiv (Y = y)$، آنگاه

$$P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{f_{X,Y}(x,y)}{f_Y(y)}, \qquad f_Y(y) \neq ۰$$

برای y ثابت، این تابع تمام خواص تابع احتمال را دارد که به آن تابع احتمال شرطی X به شرط $Y = y$ گویند و با نماد $f_{X|Y}(x|y)$ نمایش می‌دهند. به‌طور مشابه تابع احتمال شرطی Y به شرط $X = x$ نیز به‌صورت زیر تعریف می‌شود.

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f_{X,Y}(x,y)}{f_y(y)}, \qquad f_Y(y) \neq ۰ \\ f_{Y|X}(y|x) &= \frac{f_{X,Y}(x,y)}{f_X(x)}, \qquad f_X(x) \neq ۰ \end{aligned} \tag{۲۱.۱}$$

اگر X و Y دو متغیر تصادفی پیوسته باشند، تابع چگالی احتمال شرطی برای آن‌ها به همان صورت (۲۱.۱) تعریف می‌شود. همچنین برای محاسبه احتمال‌های شرطی از رابطهٔ زیر استفاده می‌کنیم.

$$P(a < X < b|Y = c) = \begin{cases} \sum_{a<X<b} f_{X|Y}(x|c) & \text{اگر } Y,X \text{ گسسته باشند} \\ \int_a^b f_{X|Y}(x|c)dx & \text{اگر } Y,X \text{ پیوسته باشند} \end{cases} \tag{۲۲.۱}$$

مثال ۱۵.۱ در مثال ۱۳.۱ تابع احتمال‌های حاشیه‌ای X و Y و تابع احتمال شرطی Y به شرط $X = ۲$ را به‌دست آورده و $P(Y \leq ۱|X = ۲)$ را محاسبه کنید.

حل : از جدول توزیع احتمالات توأم (۱۷.۱) داریم

$y \backslash x$	۰	۱	۲	۳	$f_Y(y)$
۰	۰	$\frac{۳}{۷۰}$	$\frac{۹}{۷۰}$	$\frac{۳}{۷۰}$	$\frac{۱۵}{۷۰}$
۱	$\frac{۲}{۷۰}$	$\frac{۱۸}{۷۰}$	$\frac{۱۸}{۷۰}$	$\frac{۲}{۷۰}$	$\frac{۴۰}{۷۰}$
۲	$\frac{۳}{۷۰}$	$\frac{۹}{۷۰}$	$\frac{۳}{۷۰}$	۰	$\frac{۱۵}{۷۰}$
$f_X(x)$	$\frac{۵}{۷۰}$	$\frac{۳۰}{۷۰}$	$\frac{۳۰}{۷۰}$	$\frac{۵}{۷۰}$	۱

بنابراین

y	۰	۱	۲
$f_Y(y)$	$\frac{۱۵}{۷۰}$	$\frac{۴۰}{۷۰}$	$\frac{۱۵}{۷۰}$

x	۰	۱	۲	۳
$f_X(x)$	$\frac{۵}{۷۰}$	$\frac{۳۰}{۷۰}$	$\frac{۳۰}{۷۰}$	$\frac{۵}{۷۰}$

با استفاده از

$$f_{Y|X}(y|۲) = \frac{f_{X,Y}(۲,y)}{f_X(۲)} = \frac{f_{X,Y}(۲,y)}{\frac{۳۰}{۷۰}} \quad , \qquad y = ۰,۱,۲$$

و جدول بالا داریم

y	۰	۱	۲
$f_{Y\|X}(y\|۲)$	$\frac{۹}{۳۰}$	$\frac{۱۸}{۳۰}$	$\frac{۳}{۳۰}$

بنابراین

$$P(Y \leq ۱|X = ۲) = f_{Y|X}(۰|۲) + f_{Y|X}(۱|۲) = \frac{۹}{۳۰} + \frac{۱۸}{۳۰} = \frac{۹}{۱۰}$$

مثال ۱۶.۱ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند.

$$f_{X,Y}(x,y) = \begin{cases} cx & ۰ < x < ۱, \quad ۰ < y < ۱ - x \\ ۰ & \text{سایر نقاط} \end{cases}$$

مقدار c را تعیین کنید و تابع‌های چگالی احتمال حاشیه‌ای و شرطی X و Y را به‌دست آورید.

حل : با استفاده از تعریف ۱۲.۱ داریم

$$\begin{aligned} ۱ = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{X,Y}(x,y)dydx &= \int_۰^۱\int_۰^{۱-x} cxdydx = \int_۰^۱ cx(۱-x)dx \\ &= c\left[\frac{x^۲}{۲} - \frac{x^۳}{۳}\right]_۰^۱ = c(\frac{۱}{۶}) \end{aligned}$$

بنابراین $c = ۶$. با استفاده از رابطهٔ (۲۰.۱) داریم

$$f_X(x) = \int_۰^{۱-x} ۶x dy = ۶x(۱-x), \qquad ۰ < x < ۱$$
$$f_Y(y) = \int_۰^{۱-y} ۶x dx = ۳x^۲|_۰^{۱-y} = ۳(۱-y)^۲\,, \qquad ۰ < y < ۱$$

همچنین از رابطهٔ (۲۱.۱) نتیجه می‌شود که

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{۶x}{۳(۱-y)^۲} = \frac{۲x}{(۱-y)^۲} \qquad \begin{array}{l} ۰ < y < ۱ \\ ۰ < x < ۱-y \end{array}$$

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{۶x}{۶x(۱-x)} = \frac{۱}{(۱-x)} \qquad \begin{array}{l} ۰ < x < ۱ \\ ۰ < y < ۱-x \end{array}$$

متغیرهای تصادفی مستقل در مثال ۱۴.۱ مشاهده می‌شود که

$$f_X(x) = \int_۰^۱ \frac{۱}{۴}x(۱+۳y^۲)dy = \frac{۱}{۲}x, \qquad ۰ < x < ۲$$
$$f_Y(y) = \int_۰^۲ \frac{۱}{۴}x(۱+۳y^۲)dx = \frac{۱}{۲}(۱+۳y^۲), \qquad ۰ < y < ۱$$
$$f_{X|Y}(x|y) = \frac{\frac{۱}{۴}x(۱+۳y^۲)}{\frac{۱}{۲}(۱+۳y^۲)} = \frac{۱}{۲}x = f_X(x)$$

بنابراین $f_{X|Y}(x|y)$ به y بستگی ندارد و $f_{X|Y}(x|y) = f_X(x)$. در نتیجه متغیر تصادفی Y تاثیری روی متغیر تصادفی X ندارد و در این حالت گوییم که متغیرهای تصادفی X و Y از یکدیگر مستقل هستند.

تعریف ۱۳.۱ فرض کنید متغیرهای تصادفی X و Y دارای تابع احتمال (چگالی احتمال) توأم $f_{X,Y}(x,y)$ و به‌ترتیب دارای تابع‌های احتمال (چگالی احتمال) $f_X(x)$ و $f_Y(y)$ باشند. متغیرهای تصادفی X و Y را از لحاظ آماری مستقل گویند اگر و فقط اگر

$$f_{X,Y}(x,y) = f_X(x)f_Y(y), \quad \forall\, x,y \qquad (۲۳.۱)$$

با توجه به تعریف ۱۳.۱، متغیرهای تصادفی X و Y در مثال ۱۴.۱ مستقل هستند ولی متغیرهای تصادفی X و Y در مثال‌های ۱۵.۱ و ۱۶.۱ مستقل نیستند. زیرا در مثال ۱۵.۱ داریم

$$f_{X,Y}(۱,۱) = \frac{۱۸}{۷۰} \neq \frac{۳۰}{۷۰} \times \frac{۴۰}{۷۰} = f_X(۱)f_Y(۱)$$

و در مثال ۱۶.۱ داریم

$$f_{X,Y}(x,y) = ۶x \neq [۶x(۱-x)][۳(۱-y)^۲] = f_X(x)f_Y(y)$$

توجه تمامی مباحث مربوط به توزیع احتمالات دو متغیره را می‌توان به‌راحتی به حالت چند متغیره تعمیم داد. برای مثال تابع احتمال (چگالی احتمال) توأم n متغیره تصادفی $X_۱, X_۲, \cdots, X_n$ به‌صورت $f_{X_۱,X_۲,\cdots,X_n}(x_۱, x_۲, \cdots, x_n)$ نمایش داده می‌شود و همواره $f_{X_۱,X_۲,\cdots,X_n}(x_۱, x_۲, \cdots, x_n) \geq ۰$ و

در حالت گسسته

$$۱ = \sum_{x_۱}\sum_{x_۲}\cdots\sum_{x_n} f_{X_۱,X_۲,\cdots,X_n}(x_۱, x_۲, \cdots, x_n)$$

در حالت پیوسته

$$۱ = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f_{X_۱,X_۲,\cdots,X_n}(x_۱, x_۲, \cdots, x_n)dx_۱dx_۲\cdots dx_n$$

همچنین متغیرهای تصادفی $X_۱, X_۲, \cdots, X_n$ مستقل هستند اگر و فقط اگر

$$f_{X_۱,X_۲,\cdots,X_n}(x_۱, x_۲, \cdots, x_n) = f_{X_۱}(x_۱)f_{X_۲}(x_۲)\cdots f_{X_n}(x_n)$$

و اگر علاوه بر آن متغیرهای تصادفی $X_۱, X_۲, \cdots, X_n$ هرکدام دارای تابع احتمال (چگالی احتمال) یکسان $f(x)$ باشند آنگاه

$$f_{X_۱,X_۲,\cdots,X_n}(x_۱, x_۲, \cdots, x_n) = f(x_۱)f(x_۲)\cdots f(x_n)$$

۵.۳.۱ آماره‌های ترتیبی

فرض کنید $X_۱, X_۲, \cdots, X_n$ متغیرهای تصادفی مستقل و دارای تابع توزیع یکسان $F(x)$ و تابع (چگالی) احتمال یکسان $f(x)$ باشند. اگر این متغیرهای تصادفی را از کوچک به بزرگ مرتب کنیم و قرار دهیم

$$X_{(۱)} = \min(X_۱, X_۲, \cdots, X_n)$$

$$X_{(۲)} = \text{دومین کوچک‌ترین } X_۱, X_۲, \cdots, X_n$$

$$\vdots \qquad \vdots$$

$$X_{(n)} = \max(X_۱, X_۲, \cdots, X_n)$$

دراین‌صورت $X_{(n)} \leq \cdots \leq X_{(۲)} \leq X_{(۱)}$ و $X_{(i)}$ را i امین آماره ترتیبی متغیرهای تصادفی $X_n, \cdots, X_۲, X_۱$ گویند. $X_{(n)} \leq \cdots \leq X_{(۲)} \leq X_{(۱)}$ را آماره‌های ترتیبی (یا آماره‌های مرتب) متغیرهای تصادفی $X_n, \cdots, X_۲, X_۱$ گویند. برای مثال فرض کنید یافته‌های متغیرهای تصادفی $(X_۱, X_۲, X_۳, X_۴)$ برابر $x_۱ = ۳/۲$، $x_۲ = ۱/۷$، $x_۳ = ۵$ و $x_۴ = ۰/۵$ باشد. در این‌صورت یافته‌های آماره‌های ترتیبی $(X_{(۱)}, X_{(۲)}, X_{(۳)}, X_{(۴)})$ برابر است با $x_{(۱)} = ۰/۵$، $x_{(۲)} = ۱/۷$، $x_{(۳)} = ۳/۲$ و $x_{(۴)} = ۵$.

با توجه به ساختار آماره‌های ترتیبی مشاهده می‌شود که این آماره‌ها تابعی از متغیرهای تصادفی $X_n, \cdots, X_۲, X_۱$ هستند، از یکدیگر مستقل نیستند و دارای تابع توزیع و تابع (چگالی) احتمال یکسان نیز نیستند. اگر $F_۱(x)$ و $F_n(x)$ به‌ترتیب تابع توزیع $X_{(۱)}$ و $X_{(n)}$ باشند، آنگاه با توجه به مستقل و هم توزیع بودن $X_n, \cdots, X_۲, X_۱$ داریم

$$\begin{aligned}
F_n(x) = P(X_{(n)} \leq x) &= P(X_۱ \leq x, X_۲ \leq x, \cdots, X_n \leq x) = \prod_{i=۱}^{n} P(X_i \leq x) \\
&= \prod_{i=۱}^{n} F(x) = [F(x)]^n \\
F_۱(x) = P(X_{(۱)} \leq x) &= ۱ - P(X_{(۱)} > x) \\
&= ۱ - P(X_۱ > x, X_۲ > x, \cdots, X_n > x) \\
&= ۱ - \prod_{i=۱}^{n} P(X_i > x) \\
&= ۱ - \prod_{i=۱}^{n} [۱ - F(x)] = ۱ - [۱ - F(x)]^n
\end{aligned}$$

بنابراین اگر $X_n, \cdots, X_۲, X_۱$ متغیرهای تصادفی مستقل و دارای توزیع یکسان باشند آنگاه

$$F_n(x) = [F(x)]^n, \quad F_۱(x) = ۱ - [۱ - F(x)]^n \qquad (۲۴.۱)$$

و اگر $X_n, ..., X_۲, X_۱$ متغیرهای تصادفی پیوسته باشند و $f_n(x)$ و $f_۱(x)$ به‌ترتیب تابع چگالی احتمال $X_{(n)}$ و $X_{(۱)}$ باشند، آنگاه

$$\begin{aligned}
f_n(x) &= F'_n(x) = nf(x)[F(x)]^{n-۱} \\
f_۱(x) &= F'_۱(x) = nf(x)[۱ - F(x)]^{n-۱}
\end{aligned} \qquad (۲۵.۱)$$

به همین ترتیب می‌توان نشان داد که اگر $f_r(x)$ تابع چگالی احتمال r امین آماره ترتیبی

$X_{(r)}$ باشد، آنگاه

$$f_r(x)=\frac{n!}{(r-1)!(n-r)!}\ f(x)[F(x)]^{r-1}[1-F(x)]^{n-r},\quad r=1,2,\ldots,n \tag{۲۶.۱}$$

مثال ۱۷.۱ فرض کنید $X_1,X_2,\cdots,X_n$ متغیرهای تصادفی مستقل و دارای توزیع یکسان با تابع چگالی احتمال زیر باشند

$$f_{X_i}(x)=\frac{1}{\theta}e^{-\frac{x}{\theta}},\ \ x>0,\ \theta>0,\ \ i=1,2,\cdots,n$$

و $X_{(1)}\leq X_{(2)}\leq\ldots\leq X_{(n)}$ آماره‌های ترتیبی این متغیرهای تصادفی باشند. در این‌صورت

$$\begin{aligned}
F_{X_i}(x)&=\int_{-\infty}^{x}\frac{1}{\theta}e^{-\frac{t}{\theta}}dt=1-e^{-\frac{x}{\theta}},\quad x>0\\
F_n(x)&=n\left(\frac{1}{\theta}e^{-\frac{x}{\theta}}\right)\left(1-e^{-\frac{x}{\theta}}\right)^{n-1},\quad x>0\\
F_1(x)&=n\left(\frac{1}{\theta}e^{-\frac{x}{\theta}}\right)\left(\frac{1}{\theta}e^{-\frac{x}{\theta}}\right)^{n-1}=\frac{n}{\theta}e^{-\frac{n}{\theta}x},\quad x>0\\
F_r(x)&=\frac{n!}{(r-1)!(n-r)!}\left(\frac{1}{\theta}e^{-\frac{x}{\theta}}\right)\left[1-e^{-\frac{x}{\theta}}\right]^{r-1}\left[e^{-\frac{x}{\theta}}\right]^{n-r},\ x>0
\end{aligned}$$

۴.۱ امید ریاضی

در بخش قبل مروری بر متغیرهای تصادفی و توزیع احتمال آن‌ها داشتیم. در این بخش مروری بر امید ریاضی یک متغیر تصادفی و تابع‌هایی از متغیرهای تصادفی خواهیم داشت.

تعریف ۱۴.۱ فرض کنید X متغیری تصادفی با تابع احتمال یا تابع چگالی احتمال $f_X(x)$ باشد. امید ریاضی X یا میانگین X به‌صورت زیر تعریف می‌شود.

$$\mu=E(X)=\sum_x xf_X(x) \qquad \text{اگر } X \text{ گسسته باشد}$$

$$\mu=E(X)=\int_{-\infty}^{+\infty}xf_X(x)dx \qquad \text{اگر } X \text{ پیوسته باشد} \tag{۲۷.۱}$$

درصورتی‌که مجموع یا انتگرال بالا همگرا نباشد گوییم امید ریاضی X وجود ندارد.

مثال ۱۸.۱ از جعبه‌ای محتوی ۸ لامپ که ۲ تای آن‌ها سوخته است ۳ لامپ را به تصادف انتخاب می‌کنیم. اگر X نشان دهندهٔ تعداد لامپ‌های سوخته باشد، امید ریاضی X را به‌دست آورید.

حل : تابع احتمال X به‌صورت زیر به‌دست می‌آید

$$f_X(x) = P(X = x) = \frac{\binom{۲}{x}\binom{۶}{۳-x}}{\binom{۸}{۳}}, \quad x = ۰, ۱, ۲$$

بنابراین

x	۰	۱	۲
$f_X(x)$	$\frac{۱۰}{۲۸}$	$\frac{۱۵}{۲۸}$	$\frac{۳}{۲۸}$

در نتیجه

$$\begin{aligned} \mu &= E(X) = \sum_{x=۰}^{۲} x f_X(x) \\ &= \left(۰ \times \frac{۱۰}{۲۸}\right) + \left(۱ \times \frac{۱۵}{۲۸}\right) + \left(۲ \times \frac{۳}{۲۸}\right) = \frac{۲۱}{۲۸} = \frac{۳}{۴} \end{aligned}$$

۱.۴.۱ امید ریاضی تابعی از یک یا چند متغیر تصادفی

قضیه ۱.۱ فرض کنید X متغیری تصادفی با تابع احتمال یا تابع چگالی احتمال $f_X(x)$ باشد و $g(X)$ تابعی از X باشد. امید ریاضی $g(X)$ برابر است با

$$E[g(X)] = \sum_x g(x) f_X(x) \qquad \text{اگر } X \text{ گسسته باشد}$$

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x) dx \qquad \text{اگر } X \text{ پیوسته باشد} \qquad (۲۸.۱)$$

مثال ۱۹.۱ فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال زیر باشد.

$$f_X(x) = \begin{cases} a + bx^۲ & ۰ < x < ۱ \\ ۰ & \text{سایر نقاط} \end{cases}$$

اگر $E(X) = \frac{۳}{۵}$، مقادیر a و b را تعیین کنید و $E(X^۲)$ را محاسبه کنید.

حل : برای تعیین a و b دو شرط زیر را به‌کار می‌بریم

$$\begin{aligned} ۱ &= \int_{-\infty}^{+\infty} f_X(x)dx = \int_۰^۱ (a+bx^۲)dx = ax + \frac{b}{۳}x^۳\Big|_۰^۱ = a + \frac{b}{۳} \\ \frac{۳}{۵} &= E(X) = \int_{-\infty}^{+\infty} xf_X(x)dx = \int_۰^۱ x(a+bx^۲)dx \\ &= \frac{a}{۲}x^۲ + \frac{b}{۴}x^۴\Big|_۰^۱ = \frac{a}{۲} + \frac{b}{۴} \end{aligned}$$

از حل دو معادله و دو مجهول بالا داریم $a = \frac{۳}{۵}$ و $b = \frac{۶}{۵}$. بنابراین

$$\begin{aligned} E(X^۲) &= \int_{-\infty}^{+\infty} x^۲ f_X(x)dx = \int_۰^۱ x^۲\left(\frac{۳}{۵} + \frac{۶}{۵}x^۲\right)dx \\ &= \frac{۱}{۵}\left[x^۳ + \frac{۶}{۵}x^۵\right]_۰^۱ = \frac{۱۱}{۲۵} \end{aligned}$$

تعریف ۱۵.۱ فرض کنید X و Y دو متغیر تصادفی با تابع احتمال یا تابع چگالی احتمال توأم $f_{X,Y}(x,y)$ باشند. امید ریاضی تابع $g(X,Y)$ به‌صورت زیر تعریف می‌شود.

$$E[g(X,Y)] = \sum_x \sum_y g(x,y)f_{X,Y}(x,y) \qquad \text{اگر } X,Y \text{ گسسته باشند}$$

(۲۹.۱)

$$E[g(X,Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f_{X,Y}(x,y)dxdy \qquad \text{اگر } X,Y \text{ پیوسته باشند}$$

تعریف بالا را می‌توان به‌سادگی به امید ریاضی تابعی از چند متغیر تصادفی تعمیم داد.

مثال ۲۰.۱ ظرفی حاوی ۷ مهره است که دو تا از آن‌ها با ۱، سه تا از آن‌ها با ۲ و دو تا از آن‌ها با ۳ شماره‌گذاری شده‌اند. دو مهره یک‌به‌یک و بدون جای‌گذاری از ظرف خارج می‌کنیم و متغیر تصادفی X را برابر شمارهٔ کوچک‌تر و متغیر تصادفی Y را برابر شمارهٔ بزرگ‌تر در دو مهره انتخابی در نظر می‌گیریم. $E(XY)$ را محاسبه کنید.

حل : در این مثال $S_X = \{۱,۲,۳\}$ و $S_Y = \{۱,۲,۳\}$ و همچنین

$$f_{X,Y}(۱,۱) = P(X=۱, Y=۱) = P(\{(۱,۱)\}) = \frac{۲}{۷} \times \frac{۱}{۶} = \frac{۱}{۲۱}$$

$$f_{X,Y}(۱,۲) = P(X=۱, Y=۲) = P(\{(۱,۲),(۲,۱)\}) = ۲\left(\frac{۲}{۷} \times \frac{۳}{۶}\right) = \frac{۶}{۲۱}$$

به همین ترتیب تابع احتمال توأم (X,Y) به‌صورت زیر به‌دست می‌آید.

$y \backslash x$	۱	۲	۳
۱	$\frac{۱}{۲۱}$	۰	۰
۲	$\frac{۶}{۲۱}$	$\frac{۳}{۲۱}$	۰
۳	$\frac{۴}{۲۱}$	$\frac{۶}{۲۱}$	$\frac{۱}{۲۱}$

بنابراین

$$\begin{aligned} E(X,Y) &= \sum_{x=۱}^{۳}\sum_{y=۱}^{۳} xyf_{X,Y}(x,y) \\ &= \frac{۱}{۲۱}+\frac{۱۲}{۲۱}+\frac{۱۲}{۲۱}+۰+\frac{۱۲}{۲۱}+\frac{۳۶}{۲۱}+۰+۰+\frac{۹}{۲۱}=\frac{۸۲}{۲۱} \end{aligned}$$

با استفاده از تعریف ۱۵.۱ به‌راحتی می‌توان نشان داد که

$$E[ag(X,Y)+bh(X,Y)+c]=aE[g(X,Y)]+bE[h(X,Y)]+c$$

بنابراین $E(X \pm Y)=E(X)\pm E(Y)$. همچنین اگر X و Y دو متغیر تصادفی مستقل باشند آنگاه

$$E(XY)=E(X)E(Y) \tag{۳۰.۱}$$

مثال ۲۱.۱ اگر تابع چگالی احتمال توأم (X,Y) به‌صورت زیر باشد نشان دهید که رابطهٔ (۳۰.۱) برقرار است.

$$f_{X,Y}(x,y)=\begin{cases} \frac{۱۶y}{x^{۳}} & x>۲\ ,\ ۰<y<۱ \\ ۰ & \text{سایر نقاط} \end{cases}$$

حل : توجه کنید که

$$f_X(x)=\int_۰^۱ \frac{۱۶y}{x^۳}dy=\frac{۸y^۲}{x^۳}\bigg|_۰^۱=\frac{۸}{x^۳}\ , \qquad ۰<x<۱$$

$$f_Y(y)=\int_۲^{+\infty} \frac{۱۶y}{x^۳}dx=-\frac{۸y}{x^۲}\bigg|_۲^{+\infty}=۲y\ , \qquad ۰<y<۱$$

بنابراین $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ پس X و Y مستقل هستند. همچنین

$$E(XY) = \int_0^1 \int_2^{+\infty} xy\left(\frac{16y}{x^3}\right)dxdy = \int_0^1 16y^2\left[\frac{-1}{x}\right]_2^{\infty} dy$$
$$= \int_0^1 8y^2 dy = \frac{8}{3}$$
$$E(X) = \int_2^{+\infty} x\left(\frac{8}{x^3}\right)dx = \frac{-8}{x}\Big|_2^{\infty} = 4$$
$$E(Y) = \int_0^1 y(2y)dy = \frac{2}{3}y^3\Big|_0^1 = \frac{2}{3}$$

در نتیجه $E(XY) = \frac{8}{3} = 4 \times \frac{2}{3} = E(X)E(Y)$.

۲.۴.۱ امیدهای ریاضی خاص

با استفاده (۲۸.۱) و (۲۹.۱) امید ریاضی بعضی از تابع‌های خاص را می‌توان به‌دست آورد که در این بخش آن‌ها را معرفی می‌کنیم.

گشتاورهای یک متغیر تصادفی

گشتاور r ام متغیر تصادفی X حول مبدأ به‌صورت زیر تعریف می‌شود

$$\mu_r = E(X^r) \qquad \text{گشتاور } r \text{ ام حول مبدأ} \qquad (۳۱.۱)$$

توجه کنید که $\mu_1 = E(X) = \mu$. همچنین گشتاور مرکزی مرتبهٔ r ام X (یا گشتاور مرتبهٔ r ام X حول میانگین) به‌صورت زیر تعریف می‌شود

$$\mu'_r = E[(X-\mu)^r] \qquad \text{گشتاور مرکزی مرتبه } r \text{ ام } X \qquad (۳۲.۱)$$

توجه کنید که $\mu'_0 = 1$ و $\mu'_1 = 0$ است.

واریانس گشتاور مرکزی مرتبه دوم X را واریانس X گویند و با نمادهای σ^2 یا σ_X^2 یا $\mathrm{Var}(X)$ نمایش می‌دهند

$$\sigma^2 = \sigma_X^2 = \mathrm{Var}(X) = E[(X-\mu)^2] \qquad (۳۳.۱)$$

واریانس یک متغیر تصادفی میزان پراکندگی توزیع متغیر تصادفی نسبت به میانگین آن را می‌سنجد. جذر واریانس را انحراف معیار گویند و با نماد $\sigma = \sigma_X$ نمایش می‌دهند. می‌توان

نشان داد که

$$\mathrm{Var}(X) = E(X^۲) - [E(X)]^۲$$

برای مثال در مثال ۱۹.۱، $\mathrm{Var}(X) = \frac{۱۱}{۲۵} - \left(\frac{۳}{۵}\right)^۲ = \frac{۲}{۲۵}$.

کوواریانس اگر در (۲۹.۱) قرار دهیم $g(X,Y) = (X-\mu_X)(Y-\mu_Y)$ آنگاه امید ریاضی این تابع را کوواریانس X و Y گویند و با نمادهای σ_{XY} یا $\mathrm{Cov}(X,Y)$ نمایش می‌دهند.

$$\sigma_{XY} = \mathrm{Cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] \qquad (۳۴.۱)$$

کوواریانس X و Y رابطهٔ دو متغیر تصادفی X و Y را نشان می‌دهد. اگر X و Y هم‌جهت باشند، یعنی با هم افزایش یا با هم کاهش یابند، آنگاه کوواریانس مثبت و اگر X و Y در خلاف جهت هم باشند آنگاه کوواریانس منفی است. می‌توان نشان داد که

$$\sigma_{XY} = \mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y)$$

بنابراین با توجه به رابطه (۳۰.۱) اگر X و Y مستقل باشند آنگاه $\mathrm{Cov}(X,Y) = ۰$ است. ولی عکس این مطلب برقرار نیست، یعنی اگر $\mathrm{Cov}(X,Y) = ۰$ آنگاه دلیلی ندارد که X و Y مستقل باشند.

خواص واریانس و کوواریانس اگر a، b، c و d اعداد ثابت و X و Y دو متغیر تصادفی باشند، خواص زیر را برای واریانس و کوواریانس آن‌ها می‌توان اثبات کرد.

$$\mathrm{Var}(c) = ۰, \qquad \mathrm{Var}(aX+b) = a^۲\mathrm{Var}(X)$$

$$\mathrm{Cov}(X,X) = \mathrm{Var}(X), \qquad \mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X), \qquad \mathrm{Cov}(X,c) = ۰$$

$$\mathrm{Cov}(aX+b, cY+d) = ac\mathrm{Cov}(X,Y)$$

$$\mathrm{Var}(aX+bY+c) = a^۲\mathrm{Var}(X) + b^۲\mathrm{Var}(Y) + ۲ab\mathrm{Cov}(X,Y)$$

ضریب همبستگی با توجه به خاصیت $\mathrm{Cov}(aX+b, cY+d) = ac\mathrm{Cov}(X,Y)$، مشخص است که کوواریانس بستگی به واحد اندازه‌گیری X و Y دارد. برای اینکه معیاری برای سنجش میزان رابطهٔ دو متغیر X و Y پیدا کنیم که به واحد اندازه‌گیری X و Y بستگی نداشته باشد، کوواریانس بین متغیرهای $\frac{X}{\sigma_X}$ و $\frac{Y}{\sigma_Y}$ را محاسبه می‌کنیم، یعنی

$$\mathrm{Cov}\left(\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}\right) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

این معیار را ضریب همبستگی بین متغیرهای تصادفی X و Y می‌نامند و با نمادهای ρ یا $\rho(X,Y)$ نمایش می‌دهند.

$$\rho = \rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \qquad (۳۵.۱)$$

ضریب همبستگی دو متغیر X و Y میزان رابطهٔ خطی دو متغیر X و Y را نشان می‌دهد.
خواص زیر را برای ضریب همبستگی دو متغیر تصادفی X و Y می‌توان اثبات کرد.

(الف) $\rho(aX+b, cY+d) = \rho(X,Y)$ اگر a و c هم‌علامت باشند،

(ب) $\rho(aX+b, cY+d) = -\rho(X,Y)$ اگر a و c هم‌علامت نباشند،

(پ) $-۱ \leq \rho(X,Y) \leq ۱$،

(ت) اگر $Y = aX+b$ و $a > ۰$ آنگاه $\rho = +۱$،

(ث) اگر $Y = aX+b$ و $a < ۰$ آنگاه $\rho = -۱$،

(ج) اگر X و Y مستقل باشند آنگاه $\rho = ۰$.

مثال ۲۲.۱ در مثال ۲۰.۱ ضریب همبستگی X و Y را به‌دست آورید.

حل : در مثال ۲۰.۱ داریم

x	۱	۲	۳
$f_X(x)$	$\frac{۱۱}{۲۱}$	$\frac{۹}{۲۱}$	$\frac{۱}{۲۱}$

Y	۱	۲	۳
$f_Y(y)$	$\frac{۱}{۲۱}$	$\frac{۹}{۲۱}$	$\frac{۱۱}{۲۱}$

بنابراین به‌راحتی مشاهده می‌شود که $E(X) = \frac{۳۲}{۲۱}$، $E(X^۲) = \frac{۵۶}{۲۱}$، $E(Y) = \frac{۵۲}{۲۱}$، $E(Y^۲) = \frac{۱۳۶}{۲۱}$ و در مثال ۲۰.۱ داشتیم که $E(XY) = \frac{۸۲}{۲۱}$. در نتیجه

$$\mathrm{Var}(X) = \frac{۵۶}{۲۱} - \left(\frac{۳۲}{۲۱}\right)^۲ = \frac{۱۵۲}{۴۴۱}$$

$$\mathrm{Var}(Y) = \frac{۱۳۶}{۲۱} - \left(\frac{۵۲}{۲۱}\right)^۲ = \frac{۱۵۲}{۴۴۱}$$

$$\mathrm{Cov}(X,Y) = \frac{۸۲}{۲۱} - \left(\frac{۳۲}{۲۱} \times \frac{۵۲}{۲۱}\right) = \frac{۵۸}{۴۴۱}$$

$$\rho(X,Y) = \frac{\frac{۵۸}{۴۴۱}}{\sqrt{\frac{۱۵۲}{۴۴۱} \times \frac{۱۵۲}{۴۴۱}}} = \frac{۵۸}{۱۵۲} = ۰/۳۸ > ۰$$

یعنی X و Y همبستگی ضعیف در جهت مثبت دارند.

مثال ۲۳.۱ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند. ضریب همبستگی X و Y را محاسبه کنید.

$$f_{X,Y}(x,y) = \begin{cases} e^{-y} & \cdot < x < y < +\infty \\ \cdot & \text{سایر نقاط} \end{cases}$$

حل : برای محاسبه ضریب همبستگی X و Y ابتدا تابع‌های چگالی احتمال حاشیه‌ای X و Y را به‌دست می‌آوریم.

$$f_X(x) = \int_x^{+\infty} e^{-y}dy = [-e^{-y}]_x^{+\infty} = e^{-x} \ , \quad x > \cdot$$
$$f_Y(y) = \int_\cdot^y e^{-y}dx = e^{-y}[x]_\cdot^y = ye^{-y} \ , \quad y > \cdot$$

به‌وسیله انتگرال‌گیری جزءبه‌جزء می‌توان نشان داد که برای هر عدد صحیح نامنفی n داریم $\int_\cdot^{+\infty} x^n e^{-x}dx = n!$ که بنابراین

$$E(X) = \int_\cdot^{+\infty} xe^{-x}dx = ۱! = ۱, \quad E(X^۲) = \int_\cdot^{+\infty} x^۲e^{-x}dx = ۲! = ۲$$

و در نتیجه $\text{Var}(X) = ۲ - (۱)^۲ = ۱$. همچنین

$$E(Y) = \int_\cdot^{+\infty} y^۲e^{-y}dy = ۲! = ۲, \quad E(Y^۲) = \int_\cdot^{+\infty} y^۳e^{-y}dy = ۳! = ۶$$

و در نتیجه $\text{Var}(Y) = ۶ - (۲)^۲ = ۲$. همچنین

$$\begin{aligned} E(XY) = \int_\cdot^{+\infty}\int_\cdot^y xye^{-y}dxdy &= \int_\cdot^{+\infty} ye^{-y}\left[\frac{۱}{۲}x^۲\right]_\cdot^y dy \\ &= \frac{۱}{۲}\int_\cdot^{+\infty} y^۳e^{-y}dy = \frac{۳!}{۲} = ۳ \end{aligned}$$

و در نتیجه $\text{Cov}(X,Y) = ۳ - (۱)(۲) = ۱$. بنابراین

$$\rho = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{۱}{\sqrt{(۱)(۲)}} = \frac{۱}{\sqrt{۲}} = ۰/۷۰۷$$

تابع مولد گشتاور تابع مولد گشتاور متغیر تصادفی X که آن را با نماد $M_X(t)$ نمایش می‌دهیم برای هر $t \in R$ به‌صورت زیر تعریف می‌شود

$$M_X(t) = E(e^{tX}) = \begin{cases} \sum_x e^{tx}f_X(x) & \text{اگر } X \text{ گسسته باشد} \\ \int_{-\infty}^{+\infty} e^{tx}f_X(x)dx & \text{اگر } X \text{ پیوسته باشد} \end{cases} \quad (۳۶.۱)$$

تابع $M_X(t)$ را تابع مولد گشتاور گویند زیرا همه گشتاورهای متغیرهای تصادفی X را می‌توان با مشتق‌گیری پیاپی $M_X(t)$ نسبت به t و محاسبه مقدار آن در $t = ۰$ به‌دست آورد. برای مثال

$$M'_X(۰) = \frac{d}{dt} M_X(t)|_{t=۰} = E(X)$$
$$M''_X(۰) = \frac{d^۲}{dt^۲} M_X(t)|_{t=۰} = E(X^۲)$$

به همین ترتیب

$$M^{(r)}(۰) = \frac{d^r}{dt^r} M_X(t)|_{t=۰} = E(X^r) = \mu_r \qquad (۳۷.۱)$$

مثال ۲۴.۱ فرض کنید متغیر تصادفی X دارای تابع احتمال زیر باشد. تابع مولد گشتاور X را به‌دست آورید و امید ریاضی و واریانس X را محاسبه کنید.

$$f_X(x) = \binom{n}{x} p^x (۱ - p)^{n-x}, \qquad x = ۰, ۱, ۲, ..., n$$

حل : با استفاده از بسط دو جمله‌ای داریم

$$\begin{aligned} M_X(t) = E(e^{tx}) &= \sum_{x=۰}^{n} e^{tx} \binom{n}{x} p^x (۱ - p)^{n-x} \\ &= \sum_{x=۰}^{n} \binom{n}{x} (pe^t)^x (۱ - p)^{n-x} \\ &= (pe^t + ۱ - p)^n \end{aligned}$$

همچنین

$$M'_X(t) = npe^t(pe^t + ۱ - p)^{n-۱} \Longrightarrow E(X) = M'_X(۰) = np$$
$$M''_X(t) = npe^t(pe^t + ۱ - p)^{n-۱} + n(n - ۱)p^۲ e^t (pe^t + ۱ - p)^{n-۲}$$
$$\Longrightarrow E(X^۲) = M''_x(۰) = np + n(n - ۱)p^۲$$

پس

$$\text{Var}(X) = np + n(n - ۱)p^۲ - n^۲p^۲ = np[۱ + (n - ۱)p - np] = np(۱ - p)$$

تابع مولد گشتاور یک متغیر تصادفی تابعی یکتا است، یعنی تابع مولد گشتاور بطور یکتایی توزیع یک متغیر تصادفی را مشخص می‌کند. برخی از خواص این تابع در زیر آورده شده‌اند.

(الف) اگر a و b مقادیر ثابتی باشند، آنگاه $M_{aX+b}(t) = e^{bt}M_X(at)$.

(ب) اگر X و Y متغیرهای تصادفی مستقل باشند، آنگاه

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$

(پ) اگر $X_1, X_2, \dots, X_n$ متغیرهای تصادفی مستقل و دارای توزیع یکسان باشند، آنگاه

$$M_{X_1+X_2+\dots+X_n}(t) = \prod_{i=1}^{n} M_{X_i}(t) = [M_{X_1}(t)]^n$$

(ت) اگر $X_1, X_2, \dots, X_n$ متغیرهای تصادفی مستقل و $a_1, a_2, \dots, a_n$ مقادیر ثابتی باشند، آنگاه

$$M_{a_1X_1+a_2X_2+\dots+a_nX_n}(t) = \prod_{i=1}^{n} M_{X_i}(a_it)$$

۵.۱ توزیع‌های احتمال خاص

در بخش سوم این فصل به متغیرهای تصادفی و به‌دست آوردن تابع توزیع و تابع احتمال یا تابع چگالی احتمال (توزیع احتمال) آن‌ها پرداختیم. برخی از متغیرهای تصادفی دارای الگوی خاصی هستند و می‌توان برای توزیع احتمال آن‌ها الگوی خاصی را در نظر گرفت. در این بخش برخی از توزیع‌های احتمال خاص را که کاربرد زیادی در فصل‌های آتی دارند مرور خواهیم کرد. ابتدا توزیع‌های گسسته و سپس توزیع‌های پیوسته را بررسی خواهیم کرد.

توزیع‌های احتمال گسسته

۱.۵.۱ توزیع یکنواخت گسسته

ساده‌ترین توزیع احتمال گسسته توزیعی است که درآن متغیر تصادفی X تمام مقادیرش را با احتمال‌های برابر اختیار کند. چنین توزیع احتمالی را توزیع یکنواخت گسسته گویند.

تعریف ۱۶.۱ اگر متغیر تصادفی X مقادیر $x_1, x_2, \dots, x_N$ را با احتمال‌های مساوی $\frac{1}{N}$ اختیار کند آنگاه گویند X دارای توزیع یکنواخت گسسته با پارامتر N است و آن را با نماد

$X \sim DU(N)$ نمایش می‌دهند. تابع احتمال X عبارت است از

$$\begin{aligned} f_X(x) = P(X=x) &= \frac{1}{N} \qquad x = x_1, x_2, \ldots, x_N \\ &= \frac{1}{N} I_{\{x_1,x_2,\ldots,x_N\}}(x) \end{aligned} \tag{۳۸.۱}$$

که در آن $A = \{x_1, x_2, \ldots, x_N\}$ و $I_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$ تابع نشانگر است.

مثال ۲۵.۱ در پرتاب یک تاس اگر هر شماره دارای احتمال برابر باشد و X برابر شماره روی تاس مشاهده شده باشد، آنگاه

$$X \sim DU(6), \qquad f_X(x) = \frac{1}{6} I_{\{1,2,\ldots,6\}}(x)$$

قضیه ۲.۱ اگر $X \sim DU(N)$ آنگاه

$$\sigma^2 = \mathrm{Var}(X) = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 \quad , \quad \mu = E(X) = \frac{1}{N}\sum_{i=1}^{N} x_i$$

۲.۵.۱ توزیع برنولی

آزمایش برنولی آزمایشی است که دارای دو نتیجهٔ موفقیت با احتمال p و شکست با احتمال $q = 1 - p$ است. برای مثال پرتاب یک سکه که موفقیت مشاهدهٔ شیر $(p = \frac{1}{2})$ و شکست مشاهدهٔ خط $(q = \frac{1}{2})$ و یا پرتاب یک تاس که موفقیت مشاهدهٔ عدد ۵ $(p = \frac{1}{6})$ و شکست غیر از ۵ $(q = \frac{5}{6})$ است. اگر قرار دهیم

$$X = \begin{cases} 1 & \text{اگر نتیجه آزمایش موفقیت باشد} \\ 0 & \text{اگر نتیجه آزمایش شکست باشد} \end{cases}$$

آنگاه متغیر تصادفی X را متغیر تصادفی برنولی گویند و آن را با نماد $X \sim B(1,p)$ نمایش می‌دهند. تابع احتمال توزیع برنولی عبارت است از

x	0	1
$f_X(x) = P(X=x)$	$1-p$	p

یا

$$f_X(x) = p^x(1-p)^{1-x} I_{\{0,1\}}(x) \tag{۳۹.۱}$$

قضیه ۳.۱ اگر $X \sim B(1,p)$ آنگاه

$$\mu = E(X) = p, \qquad \sigma^2 = \mathrm{Var}(X) = pq, \qquad \mu_r = E(X^r) = p$$
$$M_X(t) = E(e^{tX}) = (q + pe^t)$$

که در آن $M_X(t)$ تابع مولد گشتاور X است.

۳.۵.۱ توزیع دو جمله‌ای

آزمایش تصادفی که از انجام چند آزمایش مستقل برنولی به‌وجود می‌آید را آزمایش دو جمله‌ای گویند. برای مثال انتخاب ۵ مهره با جای‌گذاری از ظرفی که دارای ۸ مهره سفید و ۶ مهره قرمز است و موفقیت برای ما مشاهدهٔ مهرهٔ قرمز است یک آزمایش دو جمله‌ای با احتمال موفقیت $p = \frac{6}{14}$ است. به‌طورکلی یک آزمایش دو جمله‌ای دارای خواص زیر است:

۱) آزمایش دو جمله‌ای از انجام n آزمایش مستقل برنولی به‌وجود می‌آید.

۲) احتمال موفقیت p در هر آزمایش برنولی مقداری ثابت است.

تعریف ۱۷.۱ اگر متغیر تصادفی X برابر تعداد موفقیت‌ها در n آزمایش مستقل برنولی باشد، آنگاه X را یک متغیر تصادفی دو جمله‌ای با پارامترهای n و p گویند و آن را با نماد $X \sim B(n,p)$ نمایش می‌دهند. تابع احتمال X برابر است با

$$f_X(x) = \binom{n}{x} p^x q^{n-x} I_{\{0,1,\dots,n\}}(x), \quad 0 \le p \le 1 \tag{۴۰.۱}$$

قضیه ۴.۱ اگر $X \sim B(n,p)$ آنگاه

$$\begin{aligned} \mu &= E(X) = np \\ \sigma^2 &= \mathrm{Var}(X) = npq \\ M_X(t) &= E(e^{tX}) = (pe^t + q)^n \end{aligned} \tag{۴۱.۱}$$

برای اثبات به مثال ۲۴.۱ مراجعه شود.

مثال ۲۶.۱ احتمال اینکه یک بیمار از بیماری خونی خاصی شفا یابد ۰/۴ است. اگر ۱۵ نفر به چنین بیماری مبتلا شده باشند، احتمال اینکه

(الف) دقیقاً ۵ نفر زنده بمانند را بیابید.

(ب) حداقل ۱۰ نفر زنده بمانند را بیابید.

حل : اگر متغیر تصادفی X برابر تعداد بیماران شفا یافته در بین ۱۵ نفر باشد، آنگاه $X \sim B(۱۵, ۰/۴)$. بنابراین

(الف) $$P(X = ۵) = f_X(۵) = \binom{۱۵}{۵}(۰/۴)^۵(۰/۶)^{۱۰} = ۰/۱۸۵۹$$

(ب) با استفاده از جدول I پیوست پ داریم

$$P(X \geq ۱۰) = ۱ - P(X < ۱۰) = ۱ - P(X \leq ۹) = ۱ - ۰/۹۶۶۲ = ۰/۰۳۳۸$$

نکات

۱) اگر X تعداد موفقیت‌ها در n آزمایش مستقل برنولی باشد آنگاه $Y = n - X$ برابر تعداد شکست‌ها در n آزمایش مستقل برنولی است و

$$\begin{cases} X \sim B(n,p) \\ Y = n - X \sim B(n,q) \end{cases}$$

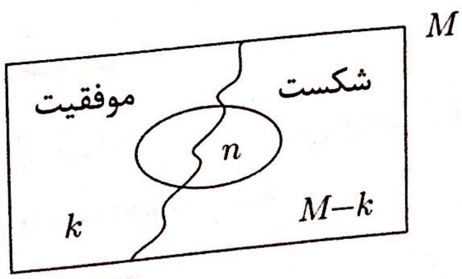
۲) اگر در n آزمایش مستقل برنولی X_i، $i = ۱, ۲, ..., n$ را نتیجهٔ آزمایش i ام در نظر بگیریم، آنگاه $X_i \sim B(۱,p)$ و اگر قرار دهیم

$$X = \sum_{i=۱}^{n} X_i = \text{تعداد موفقیت‌ها در } n \text{ آزمایش مستقل برنولی}$$

آنگاه $X = \sum_{i=۱}^{n} X_i \sim B(n,p)$.

۴.۵.۱ توزیع فوق هندسی

آزمایش تصادفی که از انجام چند آزمایش غیر مستقل برنولی به‌وجود می‌آید را یک آزمایش فوق هندسی گویند. به‌طورکلی اگر جامعهٔ مد نظر ما به دو گروه موفقیت و شکست تقسیم شده باشد که یکی k تایی و دیگری $M - k$ تایی باشد (M تعداد اعضای کل جامعه است) و بخواهیم n عضو از این جامعه M تایی بدون جای‌گذاری انتخاب کنیم، آنگاه یک آزمایش فوق هندسی داریم.



شکل ۴.۱

تعریف ۱۸.۱ اگر متغیر تصادفی X برابر تعداد موفقیت‌ها در یک آزمایش فوق هندسی باشد، آنگاه گویند X دارای توزیع فوق هندسی با پارامترهای M و k و n است و آن را با نماد $X \sim HG(M,k,n)$ نمایش می‌دهند. تابع احتمال X عبارت است از

$$f_X(x) = \frac{\binom{k}{x}\binom{M-k}{n-x}}{\binom{M}{n}} \qquad \max(۰, n-M+k) \leq x \leq \min(n,k) \qquad (۴۲.۱)$$

قضیه ۵.۱ اگر $X \sim HG(M,k,n)$ آنگاه

$$\mu = E(X) = n\frac{k}{M}$$

$$\sigma^۲ = \mathrm{Var}(X) = n\frac{k}{M}\left(۱ - \frac{k}{M}\right)\left(\frac{M-n}{M-۱}\right)$$

$$\alpha_r = E[X(X-۱)...(X-r+۱)] = r!\frac{\binom{k}{r}\binom{n}{r}}{\binom{M}{r}}$$

توجه اگر متغیر تصادفی X_i، $i = ۱,۲,...,n$ نتیجهٔ i امین آزمایش باشد آنگاه X_i ها دارای توزیع برنولی هستند ولی از یکدیگر مستقل نیستند. همچنین اگر

$$X = \sum_{i=۱}^{n} X_i = \text{تعداد موفقیت‌ها در } n \text{ آزمایش}$$

آنگاه $X \sim HG(M,k,n)$.

مثال ۲۷.۱ از بین ۷ مرد و ۴ زن یک کمیتهٔ ۳ نفری به‌طور تصادفی انتخاب می‌شود. احتمال اینکه تمام اعضای کمیته مرد باشند را بیابید. احتمال اینکه زنان کمیته بیشترین تعداد باشند را بیابید.

حل : اگر متغیر تصادفی X برابر تعداد زن‌ها در کمیته ۳ نفری باشد، آنگاه $X \sim HG(۱۱,۴,۳)$ و بنابراین

$$P(X=۰) = f_X(۰) = \frac{\binom{۴}{۰}\binom{۷}{۳}}{\binom{۱۱}{۳}} = \frac{۷}{۳۳}$$

$$P(X \geq ۲) = P(X=۲) + P(X=۳) = \frac{\binom{۴}{۲}\binom{۷}{۱} + \binom{۴}{۳}\binom{۷}{۰}}{\binom{۱۱}{۳}} = \frac{۴۶}{۱۶۵}$$

تقریب توزیع فوق هندسی به‌وسیله توزیع دو جمله‌ای اگر n نسبت به M عدد کوچکی باشد آنگاه انتخاب‌ها در مراحل مختلف احتمال‌های تقریباً یکسان دارند. یعنی انتخاب‌ها را می‌توان تقریباً با جای‌گذاری در نظر گرفت در نتیجه یک آزمایش دو جمله‌ای خواهیم داشت. بنابراین توزیع فوق هندسی را می‌توان با توزیع دو جمله‌ای با پارامترهای $p = \frac{k}{M}$ و n تقریب زد، یعنی موقعی که $M \to \infty$ و $k \to \infty$، $\frac{k}{M} = p$ باشد، داریم

$$\lim_{M\to\infty} \frac{\binom{k}{x}\binom{M-k}{n-x}}{\binom{M}{n}} = \binom{n}{x} p^x (1-p)^{n-x}$$

۵.۵.۱ توزیع دو جمله‌ای منفی

آزمایشی را در نظر بگیرید که خواص آزمایش دو جمله‌ای را داشته باشد با این تفاوت که آزمایش‌های برنولی را آن‌قدر ادامه می‌دهیم تا به‌مقدار ثابتی از موفقیت‌ها دست یابیم. چنین آزمایشی را آزمایش دو جمله‌ای منفی گویند.

تعریف ۱۹.۱ اگر متغیر تصادفی X برابر تعداد شکست قبل از مشاهده r امین موفقیت در انجام آزمایش‌های مستقل برنولی باشد، آنگاه X را یک متغیر تصادفی دو جمله‌ای منفی با پارامترهای r و p گویند و آن را با نماد $X \sim NB(r,p)$ نمایش می‌دهند. تابع احتمال X عبارت است از

$$f_X(x) = \binom{r+x-1}{x} p^r q^x I_{\{0,1,2,\dots\}}(x) \qquad (۴۳.۱)$$

قضیه ۶.۱ اگر $X \sim NB(r,p)$ آنگاه

$$\mu = E(X) = \frac{rq}{p}$$
$$\sigma^2 = \mathrm{Var}(X) = \frac{rq}{p^2}$$
$$M_X(t) = E(e^{tX}) = \left(\frac{p}{1-qe^t}\right)^r$$

مثال ۲۸.۱ شخصی پی‌درپی سه سکه را با هم پرتاب می‌کند احتمال اینکه او در بار پنجم برای دومین بار سه شیر یا سه خط به‌دست آورد را بیابید.

حل : فرض کنید $p = P(\{TTT, HHH\}) = \frac{۲}{۸} = \frac{۱}{۴}$. اگر متغیر تصادفی X برابر تعداد مشاهدهٔ سه سکهٔ غیر یکسان قبل از مشاهده دومین بار سه سکه یکسان باشد، آنگاه $X \sim NB(۲, \frac{۱}{۴})$ و در نتیجه

$$P(X = ۳) = f_X(۳) = \binom{۲ + ۳ - ۱}{۳} \left(\frac{۱}{۴}\right)^۲ \left(\frac{۳}{۴}\right)^۳ = \frac{۲۷}{۲۵۶}$$

توجه اگر در انجام آزمایش‌های برنولی با پارامتر p، متغیرهای تصادفی X و Y به‌ترتیب برابر تعداد شکست‌ها قبل از r امین موفقیت و تعداد آزمایش‌ها تا رسیدن به r موفقیت باشند، آنگاه $Y = X + r$ و Y دارای تابع احتمال زیر است.

$$\begin{aligned} f_Y(y) = p(Y = y) &= p(X = y - r) \\ &= f_X(y - r) = \binom{y - ۱}{y - r} p^r q^{y-r} I_{\{r, r+۱, \cdots\}}(y) \\ &= \binom{y - ۱}{r - ۱} p^r q^{y-r} I_{\{r, r+۱, \cdots\}}(y) \end{aligned}$$

۶.۵.۱ توزیع هندسی

اگر در توزیع دو جمله‌ای منفی $r = ۱$ باشد آنگاه یک توزیع احتمال برای تعداد شکست قبل از رسیدن به اولین موفقیت خواهیم داشت که به این آزمایش، آزمایش هندسی گویند.

تعریف ۲۰.۱ اگر متغیر تصادفی X برابر تعداد شکست قبل از اولین موفقیت در انجام آزمایش‌های مستقل برنولی باشد، آنگاه X را یک متغیر تصادفی هندسی گویند و آن را با نماد $X \sim Ge(p)$ نمایش می‌دهند. تابع احتمال X عبارت است از

$$f_X(x) = pq^x I_{\{۰, ۱, ۲, \ldots\}}(x) \qquad (۴۴.۱)$$

قضیه ۷.۱ اگر $X \sim Ge(p)$ آنگاه

$$\mu = E(X) = \frac{q}{p}, \qquad \sigma^۲ = \frac{q}{p^۲}, \qquad M_X(t) = \frac{p}{۱ - qe^t}$$

مثال ۲۹.۱ در یک خط تولید از هر ۱۰۰ قطعه به‌طور متوسط یکی خراب است. احتمال اینکه در بررسی قطعه‌ها اولین قطعهٔ خراب حداکثر در ۳ مشاهدهٔ اولیه باشد را بیابید.

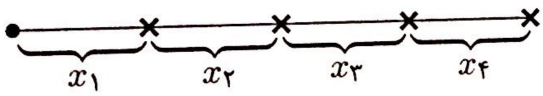
حل : اگر متغیر تصادفی X برابر تعداد قطعه‌های سالم مشاهده شده قبل از اولین قطعهٔ خراب باشد، آنگاه $X \sim Ge(۰/۰۱)$ و در نتیجه

$$\begin{aligned} P(X < ۳) = ۱ - P(X \geq ۳) &= ۱ - \sum_{x=۳}^{+\infty}(۰/۰۱)(۰/۹۹)^x \\ &= ۱ - \frac{(۰/۰۱)(۰/۹۹)^۳}{۱ - ۰/۹۹} \\ &= ۱ - (۰/۹۹)^۳ = ۰/۰۲۹۷ \end{aligned}$$

توجه اگر در یک آزمایش دو جمله‌ای منفی متغیر تصادفی X_i را برابر تعداد شکست بین $(i-۱)$ امین و i امین موفقیت در نظر بگیریم، دراین‌صورت متغیرهای تصادفی X_i از یکدیگر مستقل هستند و $X_i \sim Ge(p)$. همچنین اگر

$$X = \sum_{i=۱}^{r} X_i = \text{تعداد شکست‌ها قبل از } r \text{ امین موفقیت در آزمایش‌های مستقل برنولی}$$

آنگاه $X \sim NB(r,p)$.



شکل ۵.۱

خاصیت عدم حافظه گوییم متغیر تصادفی X دارای خاصیت عدم حافظه است هرگاه

$$P(X \geq t+s | X \geq t) = P(X \geq s)$$

یا

$$P(X \geq t+s) = P(X \geq t)P(X \geq s)$$

توزیع هندسی تنها توزیع گسسته است که دارای خاصیت عدم حافظه است.

۷.۵.۱ توزیع پواسون

آزمایشی که تعداد موفقیت در یک زمان یا ناحیه مشخص‌شده‌ای را ارائه می‌دهد آزمایش پواسون نامیده می‌شود. برای مثال تعداد تلفن‌هایی که در یک ساعت به شرکتی زده می‌شود یا تعداد غلط‌های تایپی در هر صفحه یک کتاب یا .... به‌طور کلی یک آزمایش پواسون ویژگی‌های زیر را دارد.

۱) تعداد موفقیت‌هایی که در یک فاصلهٔ زمانی یا ناحیه اتفاق می‌افتد مستقل از تعداد موفقیت‌هایی باشد که در یک فاصله زمانی دیگر یا ناحیهٔ دیگر اتفاق می‌افتد.

۲) احتمال اینکه یک موفقیت در یک فاصلهٔ زمانی کوتاه یا ناحیهٔ کوچک روی دهد، متناسب با طول فاصلهٔ زمانی یا ناحیه است و بستگی به تعداد موفقیت در خارج از فاصله زمانی یا ناحیه ندارد.

۳) احتمال روی دادن بیش از یک موفقیت در یک فاصلهٔ زمانی کوتاه یا ناحیهٔ کوچک قابل صرف‌نظر باشد.

تعریف ۲۱.۱ اگر متغیر تصادفی X برابر تعداد موفقیت‌ها در یک واحد زمانی یا ناحیهٔ مکانی معین باشد، آنگاه X را یک متغیر تصادفی پواسون گویند و اگر میانگین تعداد موفقیت‌ها در واحد زمانی یا ناحیه مکانی برابر μ باشد، گویند X دارای توزیع پواسون با پارامتر μ است و آن را با نماد $X \sim P(\mu)$ نمایش می‌دهند. تابع احتمال X عبارت است از

$$f_X(x) = \frac{e^{-\mu}\mu^x}{x!} I_{\{۰,۱,۲,\dots\}}(x) \qquad (۴۵.۱)$$

قضیه ۸.۱ اگر $X \sim P(\mu)$ آنگاه

$$E(X) = \mu, \qquad \sigma^۲ = \mathrm{Var}(X) = \mu$$

$$M_X(t) = E(e^{tX}) = e^{[\mu(e^t - ۱)]},$$

$$\alpha_r = E[X(X-۱)\cdots(X-r+۱)] = \mu^r$$

توجه اگر میانگین تعداد موفقیت در یک فاصلهٔ زمانی یا ناحیهٔ مکانی معین برابر μ باشد و X تعداد موفقیت‌ها در یک فاصلهٔ زمانی یا ناحیهٔ مکانی به‌طول t باشد، آنگاه $X \sim P(\mu t)$.

مثال ۳۰.۱ میانگین تعداد ذرات رادیواکتیو که از یک شمارنده، در یک میلیونیم ثانیه عبور می‌کند برابر ۴ است.

(الف) احتمال اینکه ۶ ذره در یک میلیونیم ثانیه از شمارنده عبور کنند را بیابید.

(ب) احتمال اینکه ۳ تا ۱۰ ذره در دو میلیونیم ثانیه از شمارنده عبور کنند را بیابید.

حل :

(الف) اگر X برابر تعداد ذرات که در یک میلیونیم ثانیه عبور می‌کنند، باشد آنگاه $X \sim P(۴)$ و

$$P(X=۶) = \frac{e^{-۴}۴^۶}{۶!} = ۰/۱۰۴۲$$

(ب) اگر Y برابر تعداد ذرات که در دو میلیونیم ثانیه عبور می‌کنند، باشد آنگاه $Y \sim P(۸)$ و با استفاده از جدول II پیوست پ داریم

$$\begin{aligned} P(۳ \leq Y \leq ۱۰) &= P(Y \leq ۱۰) - P(Y \leq ۲) \\ &= ۰/۸۱۵۹ - ۰/۰۱۳۸ = ۰/۸۰۲۱ \end{aligned}$$

تقریب توزیع دوجمله‌ای به‌وسیلهٔ توزیع پواسون

فرض کنید $X \sim B(n,p)$ و وقتی که $n \to \infty$ و $p \to ۰$ مقدار $\mu = np$ مقداری ثابت باشد دراین‌صورت

$$\lim_{n\to\infty} \binom{n}{x} p^x (۱-p)^{n-x} = \frac{e^{-\mu}\mu^x}{x!}$$

توزیع‌های احتمال پیوسته

۸.۵.۱ توزیع یکنواخت پیوسته

گوییم متغیر تصادفی X دارای توزیع یکنواخت در فاصله (a,b) است اگر دارای تابع چگالی احتمال زیر باشد

$$f_X(x) = \begin{cases} \dfrac{۱}{b-a} & a < x < b \\ ۰ & \text{سایر نقاط} \end{cases} \qquad (۴۶.۱)$$

اگر X دارای توزیع یکنواخت در فاصله (a,b) باشد آن را با نماد $X \sim U(a,b)$ نمایش می‌دهند. تابع توزیع متغیر تصادفی یکنواخت به‌صورت زیر است

$$F_X(x) = P(X \leq x) = \begin{cases} ۰ & x < a \\ \dfrac{x-a}{b-a} & a \leq x < b \\ ۱ & b \leq x \end{cases}$$

قضیه ۹.۱ اگر $X \sim U(a,b)$ آنگاه

$$\mu = E(X) = \frac{a+b}{۲}$$

$$\sigma^۲ = \mathrm{Var}(X) = \frac{(b-a)^۲}{۱۲}$$

$$M_X(t) = E(e^{tX}) = \frac{e^{bt} - e^{at}}{(b-a)t}$$

مثال ۳۱.۱ زمان لازم برای آنکه معلمی هر روز از منزلش به مدرسه بیاید متغیری تصادفی بین ۲۰ تا ۲۷ دقیقه است. اگر این معلم ساعت ۹ صبح کلاس داشته باشد و ساعت ۸ و ۳۷ دقیقه صبح منزلش را ترک کند، احتمال اینکه به‌موقع به کلاسش برسد را بیابید.

حل : اگر X زمان لازم برای رسیدن معلم از منزلش به مدرسه باشد، آنگاه $X \sim U(۲۰, ۲۷)$. بنابراین احتمال مد نظر برابر است با

$$P(X < ۲۳) = \int_{۲۰}^{۲۳} \frac{۱}{۷} dx = \frac{۳}{۷} = ۰/۴۲۸۶$$

۹.۵.۱ توزیع گاما

توزیع گاما نامش را از تابع مشهور گاما به‌دست می‌آورد که در زیر این تابع و ویژگی‌های آن را بررسی می‌کنیم.

تعریف ۲۲.۱ تابع گاما به‌صورت زیر تعریف می‌شود

$$\Gamma(\alpha) = \int_{۰}^{+\infty} x^{\alpha-۱} e^{-x} dx\ , \quad \alpha > ۰$$

قضیه ۱۰.۱ تابع گاما دارای ویژگی‌های زیر است

(الف) $\Gamma(\alpha+۱) = \alpha\Gamma(\alpha)\ , \quad \alpha > ۰$

(ب) $\Gamma(n) = (n-۱)!, \quad n = ۱, ۲, ...$

(پ) $\Gamma\left(\frac{۱}{۲}\right) = \sqrt{\pi}$

حال توزیع گاما را معرفی می‌کنیم.

تعریف ۲۳.۱ گوییم متغیر تصادفی X دارای توزیع گاما با پارامترهای α و θ است و آن را با نماد $X \sim \Gamma(\alpha, \theta)$ نمایش می‌دهند هرگاه دارای تابع چگالی احتمال زیر باشد.

$$f_X(x) = \begin{cases} \frac{۱}{\theta^\alpha \Gamma(\alpha)} x^{\alpha-۱} e^{\frac{-x}{\theta}} & x > ۰, \quad \alpha > ۰, \quad \theta > ۰ \\ ۰ & \text{سایر نقاط} \end{cases} \qquad (۴۷.۱)$$

به‌راحتی دیده می‌شود که $\int_{۰}^{+\infty} f_X(x) dx = ۱$.

قضیه ۱۱.۱ اگر $X \sim \Gamma(\alpha, \theta)$ آنگاه

$$\mu = E(X) = \alpha\theta, \quad \sigma^2 = \text{Var}(X) = \alpha\theta^2, \quad M_X(t) = (1 - \theta t)^{-\alpha} \quad t < \frac{1}{\theta}$$
$$\mu_r = E(X^r) = \frac{\Gamma(\alpha + r)}{\Gamma(\alpha)}\theta^r$$

رابطه توزیع گاما و توزیع پواسون

قضیه ۱۲.۱ اگر متغیر تصادفی X نمایانگر مدت زمان طی شده تا وقوع n امین اتفاق در یک آزمایش پواسون با میانگین μ باشد آنگاه $X \sim \Gamma\left(n, \frac{1}{\mu}\right)$.

۱۰.۵.۱ توزیع نمایی

توزیع نمایی حالت خاصی از توزیع گاما با پارامتر $\alpha = 1$ است. گوییم متغیر تصادفی X دارای توزیع نمایی با پارامتر θ است و آن را با نماد $X \sim E(\theta)$ نمایش می‌دهند، اگر دارای تابع چگالی احتمال زیر باشد.

$$f_X(x) = \begin{cases} \frac{1}{\theta}e^{\frac{-x}{\theta}} & x > 0, \theta > 0 \\ 0 & \text{سایر نقاط} \end{cases} \tag{۴۸.۱}$$

توجه معمولاً اگر متغیر تصادفی X نمایانگر طول عمر یک قطعه باشد آنگاه می‌توان X را متغیر تصادفی‌نمایی در نظر گرفت. همچنین همانند توزیع گاما اگر X نمایانگر مدت زمان طی شده تا وقوع اولین اتفاق و یا به‌طور کلی‌تر زمان بین دو اتفاق متوالی در یک آزمایش پواسون با میانگین μ باشد آنگاه $X \sim E\left(\frac{1}{\mu}\right)$ است.

مثال ۳۲.۱ به‌طور میانگین تعداد ۵ تلفن در یک ساعت به تلفن‌خانه یک شرکت زده می‌شود.

(الف) احتمال اینکه تلفن بعدی لااقل بعد از ۱۵ دقیقه زده شود را بیابید.

(ب) احتمال اینکه تلفن بعدی قبل از ۱۰ دقیقه زده شود را بیابید.

حل : اگر X برابر تعداد تلفن‌هایی که در یک ساعت به شرکت زده می‌شود و Y برابر زمان بین دو تلفن متوالی بر حسب ساعت باشد، آنگاه $X \sim P(5)$ و $Y \sim E\left(\frac{1}{5}\right)$ بنابراین

(الف)

$$P\left(Y > \frac{1}{4}\right) = \int_{\frac{1}{4}}^{+\infty} 5e^{-5y}dy = e^{-\frac{5}{4}} = ۰/۲۸۶۵$$

(ب) $P\left(Y < \frac{۱}{۶}\right) = \int_{۰}^{\frac{۱}{۶}} ۵e^{-۵y}dy = ۱ - e^{-\frac{۵}{۶}} = ۰/۵۶۵۴$

قضیه ۱۳.۱ اگر $X \sim E(\theta)$ آنگاه

$$\mu = E(X) = \theta, \quad \sigma^۲ = \mathrm{Var}(X) = \theta^۲, \quad M_X(t) = \frac{۱}{۱ - \theta t}$$

خاصیت عدم حافظه در توزیع نمایی

اگر $X \sim E(\theta)$ آنگاه

$$\begin{aligned} P(X \geq t+s | X \geq t) &= \frac{P(X \geq t+s)}{P(X \geq t)} = \frac{e^{-\frac{t+s}{\theta}}}{e^{\frac{-t}{\theta}}} \\ &= e^{\frac{-s}{\theta}} = P(X \geq s) \end{aligned}$$

یعنی X دارای خاصیت عدم حافظه است. بلعکس می‌توان نشان داد که اگر متغیر تصادفی X با تکیه‌گاه $(۰, +\infty)$ دارای خاصیت عدم حافظه باشد آنگاه X یک متغیر تصادفی نمایی است.

۱۱.۵.۱ توزیع خی‌دو (کای اسکور)

توزیع خی‌دو یک حالت خاص از توزیع گاما با پارامترهای $\alpha = \frac{n}{۲}$ و $\theta = ۲$ است که در آن n یک عدد صحیح مثبت است. گوییم متغیر تصادفی X دارای توزیع خی‌دو با n درجه آزادی است و آن را با نماد $X \sim \chi^۲_{(n)}$ نمایش می‌دهند هرگاه دارای تابع چگالی احتمال زیر باشد

$$f_X(x) = \begin{cases} \dfrac{۱}{۲^{\frac{n}{۲}}\Gamma\left(\frac{n}{۲}\right)} x^{\frac{n}{۲}-۱} e^{-\frac{۱}{۲}x} & x > ۰ \\ ۰ & \text{سایر نقاط} \end{cases} \tag{۴۹.۱}$$

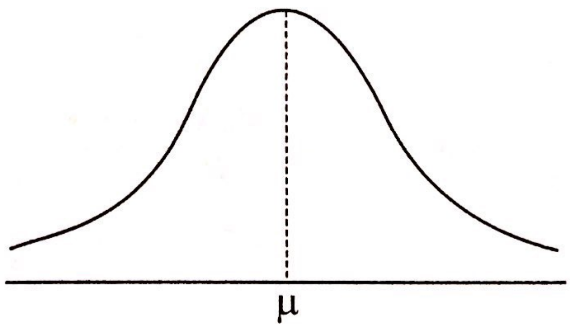
توزیع خی‌دو یکی از توزیع‌های مهم در آمار است که در استنباط آماری کاربرد فراوان دارد.

قضیه ۱۴.۱ اگر $X \sim \chi^۲_{(n)}$ آنگاه

$$\mu = E(X) = n, \quad \sigma^۲ = \mathrm{Var}(X) = ۲n, \quad M_X(t) = (۱ - ۲t)^{-\frac{n}{۲}}$$

۱۲.۵.۱ توزیع نرمال

مهم‌ترین توزیع پیوسته در آمار و احتمال توزیع نرمال است. نمودار این توزیع در شکل ۶.۱ رسم شده است و کاملاً نسبت به یک حد متوسط μ متقارن است و به آن منحنی نرمال گویند. اغلب متغیرهای تصادفی پیوسته در طبیعت و صنعت دارای نمودار توزیع مانند شکل ۶.۱ هستند. نمودار منحنی نرمال بستگی به دو پارامتر μ و $\sigma^۲$ دارد و در زیر خواهیم دید که μ میانگین توزیع و $\sigma^۲$ واریانس توزیع است.

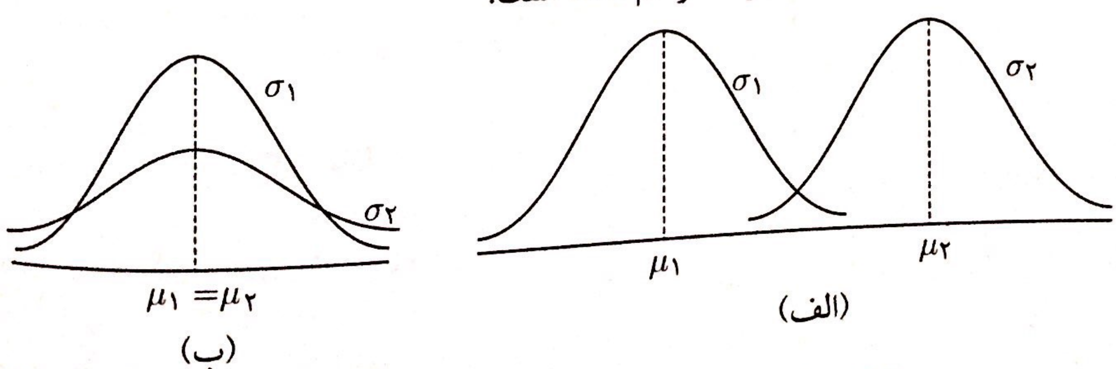


شکل ۶.۱: منحنی نرمال

تعریف ۲۴.۱ گوییم متغیر تصادفی X دارای توزیع نرمال با میانگین μ و واریانس $\sigma^۲$ است و آن را با نماد $X \sim N(\mu, \sigma^۲)$ نمایش می‌دهند، هرگاه تابع چگالی احتمال آن به‌صورت زیر باشد.

$$f_X(x) = \frac{۱}{\sqrt{۲\pi\sigma^۲}} e^{-\frac{۱}{۲\sigma^۲}(x-\mu)^۲}, \qquad \begin{array}{l} -\infty < x < +\infty \\ \mu \in R \, , \; \sigma > ۰ \end{array} \qquad (۵۰.۱)$$

منحنی نرمال کاملاً به‌وسیله مقادیر μ و $\sigma^۲$ مشخص می‌شود. با افزایش $\sigma^۲$ پراکندگی توزیع افزایش می‌یابد و با افزایش μ منحنی به سمت راست انتقال می‌یابد. در شکل ۷.۱ منحنی نرمال در حالت‌های مختلف μ و $\sigma^۲$ رسم شده است.



شکل ۷.۱: (الف) منحنی‌های نرمال با $\sigma_۱ = \sigma_۲$ و $\mu_۱ < \mu_۲$

(ب) منحنی‌های نرمال با $\sigma_۱ < \sigma_۲$ و $\mu_۱ = \mu_۲$

ویژگی‌های منحنی نرمال با استفاده از شکل تابع چگالی احتمال نرمال و همچنین نمودار آن ویژگی‌های زیر را در مورد منحنی نرمال می‌توان بیان و اثبات کرد (اثبات به‌عنوان تمرین)

۱) منحنی تنها دارای یک نقطه ماکسیمم در نقطه $X = \mu$ است.

۲) منحنی دارای دو نقطه عطف در نقاط $X = \mu \pm \sigma$ است.

۳) منحنی نسبت به خط $X = \mu$ متقارن است، یعنی $f_X(\mu + a) = f_X(\mu - a)$

۴) در دو طرف حد متوسط μ منحنی به مجانب خود یعنی محور X ها نزدیک می‌شود، یعنی

$$\lim_{x \to \pm\infty} f_X(x) = ۰$$

۵) سطح محصور بین منحنی و محور طول‌ها برابر یک واحد است، یعنی

$$\int_{-\infty}^{+\infty} f_X(x)dx = ۱$$

توزیع نرمال استاندارد

توزیع نرمالی که میانگین آن صفر و واریانس آن یک باشد را توزیع نرمال استاندارد می‌نامند و آن را با نماد $Z \sim N(۰, ۱)$ نمایش می‌دهند. تابع چگالی احتمال توزیع نرمال استاندارد را با $\phi(z)$ و تابع توزیع آن را با $\Phi(z)$ نمایش می‌دهند. بنابراین

$$\phi(z) = \frac{۱}{\sqrt{۲\pi}} e^{-\frac{۱}{۲}z^۲} \quad -\infty < z < +\infty, \quad \Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \phi(t)dt$$

در زیر خواص این توزیع و رابطه آن با توزیع‌های نرمال غیر استاندارد را بررسی می‌کنیم.

قضیه ۱۵.۱ اگر $Z \sim N(۰, ۱)$ آنگاه $E(Z) = ۰$، $\text{Var}(Z) = ۱$ و $M_Z(t) = e^{\frac{۱}{۲}t^۲}$.

قضیه ۱۶.۱ اگر $X \sim N(\mu, \sigma^۲)$ آنگاه $Z = \frac{X - \mu}{\sigma} \sim N(۰, ۱)$.

با استفاده از قضیه‌های ۱۵.۱ و ۱۶.۱ و خواص تابع مولد گشتاور می‌توان قضیه زیر را نتیجه گرفت.

قضیه ۱۷.۱ اگر $X \sim N(\mu, \sigma^۲)$ آنگاه

$$E(X) = \mu\ , \qquad \text{Var}(X) = \sigma^۲\ , \qquad M_X(t) = e^{\mu t + \frac{۱}{۲}\sigma^۲ t^۲}$$

سطح زیر منحنی نرمال

همان‌گونه که در بخش سوم مشاهده کردیم در توزیع‌های پیوسته احتمال قرار گرفتن متغیر تصادفی پیوسته X در یک فاصله (a,b) برابر سطح زیر منحنی تابع چگالی احتمال $f_X(x)$ از a تا b است که به‌صورت زیر محاسبه می‌شود

$$P(a < X < b) = \int_a^b f_X(x)dx$$

در توزیع نرمال به‌واسطهٔ پیچیده بودن فرم تابع چگالی احتمال، نمی‌توان به روش‌های معمول این انتگرال را محاسبه کرد. به این منظور در توزیع نرمال استاندارد این نوع انتگرال‌ها به‌وسیلهٔ روش‌های محاسبات عددی محاسبه و در جدولی ارائه می‌گردد و سپس با استفاده از قضیه ۱۶.۱ این مساحت‌ها در هر توزیع نرمالی محاسبه می‌گردد. برای محاسبهٔ احتمال‌ها در توزیع نرمال استاندارد جدول III در پیوست پ ارائه گردیده است که در این جدول $\Phi(z) = P(Z \leq z)$ برای مقادیر $-۳/۵ \leq z \leq ۳/۵$ محاسبه شده است. برای مثال از جدول III احتمال‌های زیر را به‌دست می آوریم.

$$\begin{aligned}
P(Z \leq ۲/۳۵) &= \Phi(۲/۳۵) = ۰/۹۹۰۶ \\
P(Z > -۱/۴) &= ۱ - P(Z \leq -۱/۴) = ۱ - \Phi(-۱/۴) \\
&= ۱ - ۰/۰۸۰۸ = ۰/۹۱۹۲ \\
P(-۰/۵۵ < Z \leq ۱/۷) &= P(Z \leq ۱/۷) - P(Z \leq -۰/۵۵) \\
&= \Phi(۱/۷) - \Phi(-۰/۵۵) \\
&= ۰/۹۵۵۴ - ۰/۲۹۱۲ = ۰/۶۶۴۲
\end{aligned}$$

حال اگر $X \sim N(\mu, \sigma^۲)$ باشد آنگاه طبق قضیه ۱۶.۱، $Z = \frac{X-\mu}{\sigma} \sim N(۰,۱)$ و بنابراین از این رابطه می‌توان برای محاسبه احتمال‌ها در هر توزیع نرمالی استفاده کرد. به مثال‌های زیر توجه کنید.

مثال ۳۳.۱ اگر $X \sim N(۱۰, ۱۶)$ باشد $P(۸ < X \leq ۱۵)$ را محاسبه کنید.

حل : در این‌جا $\mu = ۱۰$ و $\sigma^۲ = ۱۶$ است. با کم کردن مقادیر μ از سه طرف نامساوی درون پرانتز و تقسیم سه طرف بر عدد مثبت σ داریم که

$$\begin{aligned} P(۸ < X \leq ۱۵) &= P\left(\frac{۸-۱۰}{۴} < \frac{X-\mu}{\sigma} \leq \frac{۱۵-۱۰}{۴}\right) \\ &= P(-۰/۵ < Z \leq ۱/۲۵) \\ &= P(Z \leq ۱/۲۵) - P(Z \leq -۰/۵) = \Phi(۱/۲۵) - \Phi(-۰/۵) \\ &= ۰/۸۹۴۴ - ۰/۳۰۸۵ = ۰/۵۸۵۹ \end{aligned}$$

مثال ۳۴.۱ قطر داخلی پیستون‌هایی که یک کارخانه می‌سازد، توزیع نرمال با میانگین ۳۰ میلیمتر و انحراف معیار ۰/۰۲ میلیمتر دارد. احتمال اینکه قطر داخلی یک پیستون بیش از ۲۹/۹۸ باشد را بیابید.

حل : اگر متغیر تصادفی X اندازه قطر داخلی پیستون بر حسب میلیمتر باشد آنگاه $X \sim N(۳۰, ۰/۰۲^۲)$ و بنابراین

$$\begin{aligned} P(X > ۲۹/۹۸) &= P\left(\frac{X-\mu}{\sigma} > \frac{۲۹/۹۸ - ۳۰}{۰/۰۲}\right) = P(Z > -۱) \\ &= ۱ - P(Z \leq -۱) = ۱ - ۰/۱۵۸۷ = ۰/۸۴۱۳ \end{aligned}$$

مثال ۳۵.۱ یک کارگاه تولید لوله، لوله‌هایی را تولید می‌کند که طول این لوله‌ها توزیع نرمال با میانگین ۸ و انحراف معیار ۰/۵ متر دارد. مطلوبست

(الف) چند درصد لوله‌ها طولی بین ۷/۵ تا ۹ متر دارند؟

(ب) اگر در یک روز این کارگاه ۲۰۰ لوله تولید کند، چه تعداد از این لوله‌ها طولی بیش از ۹/۲ متر دارند؟

(پ) ۹۷/۵% از لوله‌های تولید شده در این کارگاه طولشان از چه مقداری کمتر است؟

(ت) اگر در یک روز بدانیم که طول لوله‌های تولید شده در این کارگاه حداقل ۷/۸ متر بوده است، چند در صد لوله‌های تولید شده در این روز طولی کمتر از ۸/۴ متر دارند.

(ث) اگر ۵ لوله را یک به یک و با جای‌گذاری انتخاب کرده و اندازه‌گیری کنیم احتمال اینکه حداکثر یک عدد از آن‌ها طولی بیش از ۹ متر داشته باشد را بیابید.

حل : اگر متغیر تصادفی X را برابر طول لوله تولید شده بر حسب متر در نظر بگیریم آنگاه $X \sim N(۸, ۰/۵^۲)$ و بنابراین

(الف) برای محاسبه درصد بایستی احتمال مد نظر را در ۱۰۰٪ ضرب کنیم، بنابراین

$$\begin{aligned} P(۷/۵ < X < ۹) &= P\left(\frac{۷/۵ - ۸}{۰/۵} < \frac{X-\mu}{\sigma} < \frac{۹ - ۸}{۰/۵}\right) = P(-۱ < Z < ۲) \\ &= \Phi(۲) - \Phi(-۱) = ۰/۹۷۷۲ - ۰/۱۵۸۷ = ۰/۸۱۸۵ \end{aligned}$$

بنابراین درصد ذکر شده برابر است با $۰/۸۱۸۵(\%۱۰۰) = \%۸۱/۸۵$.

(ب) برای محاسبه تعداد بایستی احتمال مد نظر را در تعداد ۲۰۰ لوله ضرب کنیم، بنابراین

$$\begin{aligned} P(X > ۹/۲) &= P\left(\frac{X-\mu}{\sigma} > \frac{۹/۲ - ۸}{۰/۵}\right) = P(Z > ۲/۴) \\ &= ۱ - p(Z \leq ۲/۴) = ۱ - ۰/۹۹۱۸ = ۰/۰۰۸۲ \end{aligned}$$

بنابراین

$$\text{تعداد مد نظر} = ۰/۰۰۸۲ \times ۲۰۰ = ۱/۶۴ \approx ۲$$

پس تقریباً ۲ لوله دارای طولی بیش از ۹/۲ متر هستند.

(پ) فرض کنید مقدار مد نظر a باشد بنابراین بایستی $P(X < a) = ۰/۹۷۵$ و در نتیجه

$$\begin{aligned} ۰/۹۷۵ &= P(X < a) = P\left(\frac{X-\mu}{\sigma} < \frac{a - ۸}{۰/۵}\right) \\ &= P\left(Z < \frac{a-۸}{۰/۵}\right) = \Phi\left(\frac{a-۸}{۰/۵}\right) \end{aligned}$$

بنابراین با استفاده از جدول III داریم که

$$\frac{a-۸}{۰/۵} = ۱/۹۶ \Longrightarrow a = ۸ + ۰/۵(۱/۹۶) = ۸/۹۸$$

(ت)

$$\begin{aligned} P(X < ۸/۴ | X \geq ۷/۸) &= \frac{P(۷/۸ \leq X < ۸/۴)}{P(X \geq ۷/۸)} = \frac{P(-۰/۴ \leq Z < ۰/۸)}{P(Z \geq -۰/۴)} \\ &= \frac{\Phi(۰/۸) - \Phi(-۰/۴)}{۱ - \Phi(-۰/۴)} = \frac{۰/۷۸۸۱ - ۰/۳۴۴۶}{۱ - ۰/۳۴۴۶} \\ &= ۰/۶۷۶۷ \end{aligned}$$

بنابراین درصد مد نظر برابر است با $۰/۶۷۶۷(\%۱۰۰) = \%۶۷/۶۷$.

(ث) اگر متغیر تصادفی Y برابر تعداد لوله‌هایی در بین ۵ لوله باشد که دارای طولی بیش از ۹ متر باشند آنگاه $Y \sim B(۵, p)$ که در آن

$$\begin{aligned} p &= P(X > ۹) = P\left(\frac{X-\mu}{\sigma} > \frac{۹-۸}{۰/۵}\right) = P(Z > ۲) \\ &= ۱ - P(Z \leq ۲) = ۱ - ۰/۹۷۷۲ = ۰/۰۲۲۸ \end{aligned}$$

بنابراین احتمال مد نظر برابر است با

$$\begin{aligned} P(Y \leq ۱) &= f_Y(۰) + f_Y(۱) \\ &= \binom{۵}{۰}(۰/۰۲۲۸)^۰(۰/۹۷۷۲)^۵ + \binom{۵}{۱}(۰/۰۲۲۸)^۱(۰/۹۷۷۲)^۴ \\ &= ۰/۹۹۵ \end{aligned}$$

تقریب توزیع دو جمله‌ای به‌وسیلهٔ توزیع نرمال

احتمال‌های مربوط به توزیع دو جمله‌ای را به‌راحتی می‌توان از تابع احتمال توزیع دو جمله‌ای و یا به‌وسیله جدول I محاسبه کرد. اما اگر n عدد بزرگی باشد، دیگر نمی‌توان از این جدول استفاده کرد. برای n های بزرگ می‌توان توزیع دو جمله‌ای را به‌وسیلهٔ توزیع نرمال توسط قضیه زیر تقریب زد.

قضیه ۱.۱۸ اگر X یک متغیر تصادفی دو جمله‌ای با میانگین $\mu = np$ و واریانس $\sigma^۲ = npq$ باشد آنگاه توزیع حدی $Z = \dfrac{X - np}{\sqrt{npq}}$ موقعی که $n \to +\infty$ یک توزیع نرمال استاندارد است.

معمولاً اگر $n \geq ۳۰$ باشد یا مقدار p به $\frac{۱}{۲}$ نزدیک باشد، تقریب توزیع دوجمله‌ای به‌وسیله توزیع نرمال یک تقریب مناسب است. همچنین چون توزیع دو جمله‌ای یک توزیع گسسته و توزیع نرمال یک توزیع پیوسته است، در موقع تقریب زدن بایستی در محاسبه احتمال‌ها از عمل تصحیح پیوستگی به‌صورت زیر استفاده کنیم.

$$P(X = K) = P\left(K - \frac{۱}{۲} < X < K + \frac{۱}{۲}\right), \quad P(X \leq K) = P\left(X < K + \frac{۱}{۲}\right)$$

مثال ۱.۳۶ اگر $X \sim B(۲۰, ۰/۶)$ باشد، احتمال $P(۱۰ \leq X \leq ۱۴)$ را به‌صورت دقیق و تقریبی محاسبه کنید.

حل : با استفاده از جدول I مقدار دقیق برابر است با

$$P(۱۰ \leq X \leq ۱۴) = P(X \leq ۱۴) - P(X \leq ۹) = ۰/۸۷۴۴ - ۰/۱۲۷۵ = ۰/۷۴۶۹$$

همچنین چون $\mu = np = ۲۰(۰/۶) = ۱۲$ و $\sigma^۲ = npq = ۲۰(۰/۶)(۰/۴) = ۴/۸$ پس مقدار تقریبی با استفاده از عمل تصحیح پیوستگی به‌صورت زیر به‌دست می‌آید

$$\begin{aligned} P(۱۰ \leq X \leq ۱۴) &= P(۹/۵ < X < ۱۴/۵) \\ &= P\left(\frac{۹/۵ - ۱۲}{\sqrt{۴/۸}} < \frac{X - np}{\sqrt{npq}} < \frac{۱۴/۵ - ۱۲}{\sqrt{۴/۸}}\right) \\ &\approx P(-۱/۱۴ < Z < ۱/۱۴) = \Phi(۱/۱۴) - \Phi(-۱/۱۴) \\ &= ۰/۸۷۲۹ - ۰/۱۲۷۱ = ۰/۷۴۵۸ \end{aligned}$$

که اختلاف این دو مقدار ۰/۰۰۱۱ است.

مثال ۳۷.۱ احتمال اینکه یک قطعه الکترونیکی در کمتر از ۱۰۰۰ ساعت استفاده مداوم از کار بیفتد، برابر ۰/۲۵ است. احتمال اینکه در بین ۲۰۰ عدد از این قطعه‌ها کمتر از ۴۵ قطعه در کمتر از ۱۰۰۰ ساعت استفاده مداوم از کار بیفتند را بیابید.

حل : اگر X تعداد قطعه‌های الکترونیکی در بین ۲۰۰ قطعه باشد که در کمتر از ۱۰۰۰ ساعت استفاده مداوم از کار می‌افتند، آنگاه $X \sim B(۲۰۰, ۰/۲۵)$ و بنابراین

$$\begin{aligned} P(X < ۴۵) &= P(X < ۴۴/۵) = P\left(\frac{X - np}{\sqrt{npq}} < \frac{۴۴/۵ - ۵۰}{\sqrt{۳۷/۵}}\right) \\ &\approx P(Z < -۰/۹) = ۰/۱۸۴۱ \end{aligned}$$

توزیع مجموع متغیرهای تصادفی نرمال

در برخی مسئله‌ها نیاز به محاسبهٔ توزیع مجموعی از چند متغیر تصادفی نرمال مستقل داریم که این توزیع در قضیه زیر آورده شده است.

قضیه ۱۹.۱ اگر $X_۱, X_۲, \ldots, X_n$ متغیرهای تصادفی مستقل باشند و

$$X_i \sim N(\mu_i, \sigma_i^۲), \qquad i = ۱, ۲, \ldots, n$$

و قرار دهیم $Y = a_1X_1 + a_2X_2 + ... + a_nX_n = \sum_{i=1}^{n} a_iX_i$ دراین‌صورت

$$Y = \sum_{i=1}^{n} a_iX_i \sim N(\sum_{i=1}^{n} a_i\mu_i, \sum_{i=1}^{n} a_i^2\sigma_i^2) \quad (51.1)$$

برهان با استفاده از خواص تابع مولد گشتاور و قضیه ۱۷.۱ نتیجه به‌راحتی حاصل می‌شود.

مثال ۳۸.۱ نمره یک درس در امتحان میان‌ترم دارای میانگین ۵ و انحراف معیار ۳ و در امتحان پایان‌ترم دارای میانگین ۶ و انحراف معیار ۴ است. فرض کنید نمره‌های دو امتحان از یکدیگر مستقل بوده و هر یک توزیع نرمال داشته باشند. شخصی در این درس قبول می‌شود که ۲ برابر نمره میان‌ترم او بعلاوه ۲ برابر نمره پایان‌ترم او از ۱۸ کمتر نباشد. اگر در این درس ۵۰ نفر ثبت نام کرده باشند، چند نفر آن‌ها رد می‌شوند؟

حل : اگر متغیرهای تصادفی X_1 و X_2 به‌ترتیب نمره‌های امتحان شخص در میان‌ترم و پایان‌ترم باشد و Y نمره کل او باشد آنگاه

$$X_1 \sim N(5, 3^2), \qquad X_2 \sim N(6, 4^2)$$
$$Y = 2X_1 + 2X_2 \sim N(2(5) + 2(6), 2^2(3^2) + 2^2(4^2))$$

بنابراین $Y \sim N(22, 100)$ و در نتیجه

$$P(Y < 18) = P(\frac{Y-\mu}{\sigma} < \frac{18-22}{10}) = P(Z < -0/4) = 0/3446$$

بنابراین تعدادی که در این درس رد می‌شوند برابر است با $17 \approx 17/23 = 50 \times 0/3446$.

۶.۱ تمرین‌ها

۱) برای هر یک از آزمایش‌های زیر فضای نمونه را بنویسید.

(الف) پرتاب یک سکه تا مشاهدهٔ اولین شیر.

(ب) پرتاب یک تاس تا مشاهدهٔ خال مضرب ۳.

(پ) پرتاب دو تاس و مشاهدهٔ مجموع خال‌ها.

(ت) انتخاب یک نقطه به تصادف در داخل دایره‌ای به مرکز محور مختصات و شعاع r.

(ث) انتخاب یک نقطه به تصادف در فاصله $[۰, ۱]$.

(ج) انتخاب دو مهره به تصادف، یک‌به‌یک و بدون جای‌گذاری از جعبه‌ای که شامل ۴ مهره سفید و ۲ مهره سیاه است.

(چ) پرتاب دو تاس و مشاهدهٔ تفاضل خال‌ها.

۲) برای هر یک از فضاهای نمونه سوال ۱، یک پیشامد مناسب تعریف کنید و به‌صورت یک مجموعه نمایش دهید.

۳) فضای نمونهٔ مربوط به آزمایش قرار دادن سه کتاب متفاوت در یک قفسه به‌صورت تصادفی را تعریف کنید. اگر دو جلد از این کتاب‌ها لغت‌نامه باشند، پیشامد قرار گرفتن این دو جلد پهلوی هم و به‌ترتیب شماره (یعنی اول جلد ۱ بعد جلد ۲) را تعریف کنید.

۴) اتومبیلی که هر روز مسافران هواپیمایی را به سه هتل مختلف می‌رساند فرودگاه را با دو مسافر ترک کرده است. فضای نمونهٔ نقاط توقف را تعیین کرده و پیشامد اینکه هر دو مسافر در یک هتل پیاده شوند، مشخص کنید.

۵) اگر A و B دو پیشامد ناسازگار باشند و $P(A) = ۰/۳$ و $P(B) = ۰/۴$ مطلوب است محاسبهٔ

$$P(A \cup B), \qquad P(B - A), \qquad P(B' - A')$$

۶) احتمال اینکه به بیماری که به دکتر مراجعه کرده، قرص داده شود ۰/۳۷ و احتمال اینکه به وی آمپول داده شود ۰/۲۸ است و احتمال اینکه به وی هم قرص و هم آمپول داده شود ۰/۱۵ است. احتمال اینکه به وی قرص یا آمپول و یا هر دو داده شود را بیابید.

۷) در یک بررسی از اهداکنندگان خون، ۲۲۵ نفر در گروه O و ۲۷۵ نفر در گروه غیر O طبقه‌بندی شده‌اند. احتمال تقریبی فردی که دارای گروه خونی O است چقدر است؟

۸) اگر یک نفر به تصادف انتخاب شود، احتمال اینکه

(الف) روز تولد وی، اول آبان ماه باشد چقدر است؟

(ب) روز تولد وی، یکی از روزهای ماه آبان باشد چقدر است؟

۹) جدول زیر تعداد ۲۰ نفر را برحسب جنس و پیری و جوانی نشان می‌دهد. اگر این عده بخواهند یک کمیته سه نفری از بین خود انتخاب کنند،

(الف) احتمال اینکه هر سه نفر انتخاب شده زن باشند چیست؟
(ب) احتمال اینکه هر سه نفر انتخاب شده مرد باشند چیست؟
(پ) احتمال اینکه هر سه نفر انتخاب شده از یک جنس باشند چیست؟
(ت) احتمال اینکه هر سه نفر انتخابی جوان باشند چیست؟
(ث) احتمال اینکه هر سه نفر انتخابی زن یا جوان باشند چیست؟

جنس \ پیری و جوانی	جوان	پیر
مرد	۷	۴
زن	۵	۴

۱۰) ۱۲ بسته محتوی جوایز در دست داریم که قرار است به دانش‌آموزی یکی از این ۱۲ بسته جایزه داده شود. دانش‌آموز نمی‌داند که در ۳ بسته کیف بغلی، در ۵ بسته دفترچه و در ۴ بسته خودکار وجود دارد.
(الف) اگر دانش‌آموز یک بسته را به‌تصادف انتخاب کند، احتمال اینکه جایزه او کیف بغلی باشد را بیابید.
(ب) اگر دانش‌آموز دو بسته را به تصادف انتخاب کند، احتمال اینکه جایزه او هر دو کیف بغلی باشد را بیابید.

۱۱) ده مهرهٔ غیرمتمایز را به چند طریق می‌توانید در چهار جعبه متمایز قرار دهید درحالی که در هر جعبه لااقل یک مهره قرار گیرد؟

۱۲) یک خانواده سه نفری، پنجشنبه یا جمعه متولد شده‌اند. احتمال اینکه هم پنجشنبه و هم جمعه جشن تولد داشته باشند را بیابید.

۱۳) با فرض $r \leq n$ ثابت کنید که

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1} \;, \qquad \binom{n}{r} = \binom{n}{n-r}$$

۱۴) بیست اتومبیل در مسابقه‌ای شرکت می‌کنند. هشت مرسدس بنز، هفت پژو و بقیه شورلت می‌باشند. هر گاه فقط کارخانه سازنده اتومبیل در نظر گرفته شود، مطلوب است.
(الف) احتمال اینکه یکی از اتومبیل‌های مرسدس بنز اول شود را بیابید.
(ب) احتمال اینکه اتومبیل‌های مرسدس بنز مقام‌های اول و دوم را کسب کنند،

بیابید.

۱۵) یک آسانسور از طبقه همکف با ۸ نفر حرکت می‌کند. به‌طوری‌که در طبقه ششم هیچ‌کس در آسانسور باقی‌نمی‌ماند. به چند روش این افراد می‌توانسته‌اند از آسانسور پیاده شوند؟ اگر افراد شامل ۵ مرد و ۳ زن بوده باشند، جواب چیست؟

۱۶) نقطه‌ای تصادفی بین ۰ و ۱ روی محور x ها در صفحه (x, y) انتخاب می‌کنیم سپس دایره‌ای به مرکز مبدأ مختصات رسم می‌کنیم که برآن نقطه بگذرد. احتمال اینکه مساحت دایره رسم شده کمتر از $\frac{\pi}{۲}$ باشد را بیابید.

۱۷) یک خط‌کش ۳۰ سانتی‌متری به تصادف در یک نقطه در امتداد طولش به دو قطعه شکسته شده است. احتمال اینکه قطعهٔ بلندتر لااقل دو برابر قطعه کوتاه‌تر باشد را بیابید.

۱۸) ثابت کنید روابط زیر برقرار هستند:

(الف) $\binom{n}{r} = \binom{n-۱}{r-۱} + \binom{n-۱}{r}$ (ب) $\binom{۲n}{r} = \sum_{r=۰}^{n} \binom{n}{r}^۲$

۱۹) روابط زیر را ثابت کنید.

(الف) $k\binom{n}{k} = (n-k+۱)\binom{n}{k-۱} = n\binom{n-۱}{k-۱}$

(ب) $\sum_{k=۱}^{n} k\binom{n}{k} = n۲^{n-۱}$

(پ) $\sum_{k=۱}^{n} k^۲\binom{n}{k} = n(n+۱)۲^{n-۲}$

۲۰) احتمال اینکه جراحی روش خاصی در جراحی را انتخاب کند ۰/۴ و احتمال اینکه همکارش نیز همان روش را انتخاب کند ۰/۵ و احتمال اینکه جراح روش همکارش را انتخاب کند ۰/۷ است. مطلوب است:

(الف) احتمال اینکه هر دو یک روش را انتخاب کنند را بیابید.

(ب) احتمال اینکه همکار روشی را انتخاب کند که جراح انتخاب کرده است را بیابید.

(پ) احتمال اینکه روش خاصی در جراحی حداقل به‌دست یکی از این دو انتخاب شود را بیابید.

۲۱) سه سکه وجود دارد که سکه اول یک سکه سالم است، سکه دوم دو شیری (هر دو طرف آن شیر) است و سکه سوم شانس مشاهده شیر در آن سه برابر خط است. یکی از سکه‌ها را به‌تصادف انتخاب می‌کنیم. احتمال اینکه در پرتاب یک مرتبه این

سکه شیر مشاهده شود را بیابید.

۲۲) فروشگاهی سه شعبه A، B و C دارد که به‌ترتیب ۵۰، ۷۵ و ۱۰۰ کارمند دارند و به‌ترتیب ۵۰، ۶۰ و ۷۰ درصد کارمندان آن شعب زن هستند. فرض کنید استعفا در میان تمام کارمندان یکسان است.

(الف) اگر از هر شعبه یک نفر استعفا دهد، احتمال آنکه هر سه نفر زن باشند را بیابید.

(ب) اگر از این فروشگاه یک نفر استعفا دهد. احتمال آنکه زن باشد را بیابید.

۲۳) مقدار c را طوری به‌دست آورید که هر یک از تابع‌های زیر یک تابع احتمال برای متغیر تصادفی X باشند.

(الف) $f_X(x) = c(x^۲ + ۴) \qquad x = ۰, ۱, ۲, ۳$

(ب) $f_X(x) = c\binom{۲}{x}\binom{۳}{۳ - x} \qquad x = ۰, ۱, ۲$

۲۴) یک شرکت سرمایه‌گذاری پیشنهاد نوعی سرمایه‌گذاری می‌کند که درصد سود را از کل سرمایه توسط تابع توزیع زیر طبق سال‌های سرمایه‌گذاری محاسبه می‌کند.

$$F_T(t) = \begin{cases} ۰ & t < ۱ \\ \frac{۱}{۴} & ۱ \leq t < ۳ \\ \frac{۱}{۲} & ۳ \leq t < ۵ \\ \frac{۲}{۳} & ۵ \leq t < ۷ \\ ۱ & t \geq ۷ \end{cases}$$

برای هر یک از حالت‌های زیر چند درصد سود محاسبه می‌شود.

(الف) اگر سرمایه‌گذاری به مدت ۵ سال شود.

(ب) اگر سرمایه‌گذاری از ۳ سال بیشتر شود.

(پ) اگر سرمایه‌گذاری بین ۱/۴ سال و ۶ سال شود.

۲۵) اگر تابع توزیع یک متغیر تصادفی X به‌صورت زیر باشد

$$F_X(x) = \begin{cases} ۰ & x < ۲ \\ c - \frac{۴}{x^۲} & x \geq ۲ \end{cases}$$

(الف) مقدار c را تعیین کنید و تابع چگالی احتمال X را به‌دست آورید.

(ب) $P(X < ۳)$ و $P(۴ < X \leq ۵)$ را محاسبه کنید.

۲۶) تابع احتمال توأم متغیرهای تصادفی X و Y به‌صورت زیر است

$$f_{X,Y}(x,y) = k(x+y) \qquad x = ۰,۱,۲,۳\ ,\ y = ۰,۱,۲$$

(الف) مقدار k را تعیین کنید.

(ب) احتمال‌های $P(X+Y=۴)$ و $P(X \leq ۲, Y=۱)$ را به‌دست آورید.

۲۷) فرض کنید X زمان واکنش نسبت به یک محرک برحسب ثانیه و Y دمایی برحسب فارنهایت باشد که یک محرک شروع به واکنش می‌کند. تابع چگالی احتمال توأم X و Y به‌صورت زیر است

$$f_{X,Y}(x,y) = \begin{cases} ۴xy & ۰ < x < ۱\ ,\ ۰ < y < ۱ \\ ۰ & \text{سایر نقاط} \end{cases}$$

(الف) $P\left(۰ \leq X \leq ۱۲, \frac{۱}{۴} \leq Y \leq \frac{۱}{۲}\right)$ را پیدا کنید.

(ب) $P(X < Y)$ را به‌دست آورید.

(پ) توزیع کناری X را محاسبه کنید.

(ت) $f_{X|Y}(x|y)$ را محاسبه کنید.

۲۸) فرض کنید X و Y دو متغیر تصادفی مستقل با تابع‌های توزیع زیر باشند

$$F_X(x) = \begin{cases} ۰ & x < ۰ \\ \frac{x}{۲} & ۰ \leq x < ۱ \\ ۱ & ۱ \leq x \end{cases} \quad , \quad F_Y(y) = \begin{cases} ۰ & y < ۰ \\ \frac{y^۴}{۱۶} & ۰ \leq y < ۲ \\ ۱ & ۲ \leq y \end{cases}$$

مطلوب است محاسبه $P\left(XY > \frac{۱}{۲}\right)$ و $P\left(۰ < X < \frac{۱}{۲} \middle| Y = ۱\right)$.

۲۹) در یک نمونه‌گیری، متغیر تصادفی N را تعداد مشاهده‌ها و متغیر تصادفی R را تعداد موفقیت‌ها در نظر می‌گیریم. این دو متغیر تصادفی دارای تابع احتمال توأم زیر هستند. مقدارهای $P(R=۲, N=۵)$، $P(R \leq ۳, N=۵)$ و $P(R \leq ۱|N=۳)$ را محاسبه کنید.

$$P_{R,N}(r,n) = \frac{e^{-۲}}{r!(n-r)!} \ , \qquad r = ۰,۱,۲,\cdots,n \qquad n = ۰,۱,۲,\cdots$$

۳۰) فرض کنید B و C دو متغیر تصادفی مستقل و هم‌توزیع با تابع چگالی احتمال زیر

باشند

$$f(x) = \begin{cases} \frac{۳x^۲}{۳۶} & ۱ < x < ۳ \\ ۰ & \text{سایر مقادیر} \end{cases}$$

احتمال اینکه معادلهٔ درجه دوم $X^۲ + BX + C = ۰$ ریشه حقیقی داشته باشد چقدر است؟

۳۱) X و Y دو متغیر تصادفی پیوسته مستقل هستند. اگر

$$P(X > ۰) = P(Y > ۰) = \frac{۱}{۲}$$

$P(XY > ۰)$ را محاسبه کنید.

۳۲) اگر متغیرهای تصادفی $X_۱, \cdots, X_n$ مستقل از هم و دارای توزیع یکسان با تابع چگالی احتمال زیر باشد، تابع چگالی احتمال هرکدام از متغیرهای تصادفی $X_{(۱)}$ و $X_{(n)}$ را پیدا کنید.

$$f_X(x) = e^{\theta - x}, \quad x \geq \theta, \ -\infty < \theta < \infty$$

۳۳) فرض کنید برای متغیرهای تصادفی X و Y داشته باشیم که

$$f_X(x) = \frac{e^{-\mu}\mu^x}{x!} \qquad x = ۰, ۱, ۲, \cdots$$

$$f_{Y|X}(y|x) = \binom{x}{y} p^y (۱ - p)^{x-y} \qquad y = ۰, ۱, \cdots, x$$

که در آن $۰ < \mu$ و $۱ > p > ۰$ مقادیر ثابتی هستند. نشان دهید که تابع احتمال حاشیه‌ای Y برابر است با

$$f_Y(y) = \frac{e^{-\mu p}(\mu p)^y}{y!} \qquad y = ۰, ۱, ۲, \cdots$$

۳۴) متغیرهای تصادفی پیوسته X و Y با تابع چگالی توأم زیر مفروض است

$$f_{X,Y}(x,y) = \begin{cases} \frac{x(۱ + ۳y^۲)}{۴} & ۰ < x < ۲, \ ۰ < y < ۱ \\ ۰ & \text{در غیر این‌صورت} \end{cases}$$

$E\left(\frac{Y}{X}\right)$ را محاسبه کنید. آیا X و Y همبسته‌اند؟

۳۵) واریانس و انحراف معیار متغیر تصادفی X که دارای تابع احتمال زیر است را حساب کنید

$$f_X(x) = \begin{cases} \dfrac{|x-۳|+۱}{۲۸} & x = -۳, -۲, -۱, ۰, ۱, ۲, ۳ \\ ۰ & \text{در غیر این‌صورت} \end{cases}$$

۳۶) فرض کنید X و Y دو متغیر تصادفی مستقل از هم هستند که فقط مقدارهای $-۱, ۰, ۲$ را اختیار می‌کنند به‌گونه‌ای که

$$P(X=-۱) = P(X=۲) = \frac{۱}{۳} \quad , \quad P(Y=-۱) = P(Y=۰) = \frac{۱}{۴}$$

اگر $Z = ۵X + ۴Y$، مقدار $E(Z)$ و $\mathrm{Var}(Z)$ را به‌دست آورید.

۳۷) فرض کنید متغیرهای تصادفی X و Y دارای تابع احتمال توأم زیر هستند. مقدار $\mathrm{Cov}(X,Y)$ را محاسبه کنید.

y \ x	-۱	۰	۱
-۱	$\frac{۱}{۶}$	$\frac{۱}{۶}$	$\frac{۱}{۶}$
۱	$\frac{۱}{۶}$	$\frac{۱}{۶}$	$\frac{۱}{۶}$

۳۸) ثابت کنید که

$$\mathrm{Var}(X-Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) - ۲\mathrm{Cov}(X,Y)$$

۳۹) فرض کنید X و Y دو متغیر تصادفی مستقل باشند.

(الف) نشان دهید $X-Y$ و $X+Y$ ناهم‌بسته‌اند، اگر و تنها اگر

$$\mathrm{Var}(X) = \mathrm{Var}(Y)$$

(ب) نشان دهید $\mathrm{Cov}(X, XY) = E(Y)\mathrm{Var}(X)$.

۴۰) متغیر تصادفی X با تابع احتمال زیر مفروض است

$$f_X(x) = \frac{e^{-\theta}\theta^x}{x!} \qquad x = ۰, ۱, ۲, \cdots$$

تابع مولد گشتاور X را به‌دست آورید و با استفاده از آن امیدریاضی و واریانس آن را به‌دست آورید.

(۴۱) فرض کنید X و Y متغیرهای تصادفی مستقل با تابع چگالی‌های احتمال زیر باشند

$$f_X(x) = \begin{cases} \frac{8}{x^3} & x < 2 \\ 0 & \text{سایر نقاط} \end{cases} \quad , \quad f_Y(y) = \begin{cases} 2y & 0 < y < 1 \\ 0 & \text{سایر نقاط} \end{cases}$$

$M_{X+Y}(t)$ و $E(XY)$ را محاسبه کنید.

(۴۲) از ده کارمندی که به شماره‌های ۱ تا ۱۰ شماره‌گذاری کرده‌ایم یک کارمند را به تصادف انتخاب می‌کنیم. اگر X شماره کارمند انتخابی باشد، احتمال اینکه شماره کارمند کمتر از ۴ باشد را بیابید. میانگین و واریانس X را به‌دست آورید.

(۴۳) فرض کنید موتورهای هواپیما به‌طور مستقل از یکدیگر با احتمال $\frac{1}{5}$ خراب می‌شوند و اگر حداقل نصف از موتورهای هواپیما کار کند، هواپیما سالم به زمین خواهد نشست. کدام یک از هواپیماهای چهار موتوره یا دو موتوره احتمال بیشتری برای یک پرواز موفق دارند.

(۴۴) فرض کنید $X_1, \cdots, X_n$ متغیرهای تصادفی مستقل و هرکدام دارای توزیع برنولی با پارامتر p باشند. مقدار $P(\max(X_1, \cdots, X_n) = 1)$ را تعیین کنید.

(۴۵) یک امتحان تستی ۵ جوابی شامل ۱۰ سوال است. برای قبولی در امتحان نمره ۵۰٪ مورد نیاز است (نیاز به ۵ جواب صحیح از ۱۰ سوال است). احتمال قبولی در امتحان چقدر است. اگر

(الف) شما در این امتحان بدون اطلاعات و برحسب تصادف به سؤال‌ها جواب دهید.

(ب) شما به اندازهٔ کافی مطالعه داشته‌اید و برای هر سوال ۳ جواب را کنار گذاشته، آنگاه بین دو جواب باقیمانده با حدس و گمان و به‌تصادف جواب دهید.

(۴۶) از ۱۰۰۰ نفر ساکن در یک منطقه، ۶۰۰ نفر خودروهای ساخت داخل دارند. اگر ۲۰ نفر از افراد این منطقه انتخاب کنیم. احتمال اینکه حداقل ۱۲ نفر اتومبیل ساخت خارج داشته باشند را بیابید.

(۴۷) فروشگاهی ۵۰ دستگاه ضبط صوت را از تولید کننده‌ای دریافت می‌کند. چهار دستگاه به‌تصادف انتخاب شده و آزمایش می‌شوند. اگر هر چهار دستگاه سالم باشند فروشگاه محموله را قبول می‌کند و در غیر این صورت ۵۰ دستگاه را پس می‌دهد. فرض کنید محموله حاوی ۵ دستگاه نقص‌دار است.

(الف) احتمال اینکه محموله را فروشگاه بپذیرد بیابید.

(ب) احتمال اینکه دو دستگاه آزمایش شدهٔ اول نقص‌دار و دو دستگاه بعدی بی‌نقص

باشد را بیابید.

۴۸) شخصی می‌خواهد بسته‌ای شامل ۱۵۰ قطعه الکترونیکی را بفروشد. او نمی‌داند ۳۰ عدد از این قطعه‌ها معیوب هستند. خریدار در صورتی حاضر به خرید این محصول است که در یک نمونهٔ ۱۰ تایی کمتر از ۲ قطعه معیوب وجود داشته باشد. فروشنده محصولات را به مشتریان مختلف نشان می‌دهد. احتمال اینکه در سومین تلاش، موفق به فروش محصولات شود چقدر است؟

۴۹) در پرتاب متوالی یک تاس،

(الف) احتمال اینکه سومین شش، بعد از هفتمین آزمایش و قبل از دهمین آزمایش مشاهده شود را بیابید.

(ب) احتمال اینکه در دهمین پرتاب اولین عدد مضرب ۳ مشاهده شود را بیابید.

۵۰) در یک اتوبان در هر شبانه روز به‌طور میانگین ۳ تصادف اتومبیل روی می‌دهد.

(الف) اگر بدانیم در یک شبانه روز حداقل ۲ تصادف روی داده است، احتمال اینکه در همان شبانه روز حداکثر ۳ تصادف روی داده باشد را بیابید.

(ب) احتمال اینکه زمان بین دو تصادف متوالی حداقل ۸ ساعت باشد را بیابید.

۵۱) از ساعت ۷ الی ۸ صبح هر روز به‌طور متوسط ۶ اتومبیل به یک پمپ بنزین وارد می‌شود. در ساعت ۷:۳۰ یک روز خاص مشاهده شده که فقط یک اتومبیل وارد پمپ بنزین شده است.

(الف) احتمال اینکه حداکثر تا ساعت ۷:۴۰ همان روز اتومبیل بعدی وارد شود را بیابید.

(ب) احتمال اینکه تا ساعت ۸ همان روز هیچ اتومبیل دیگری وارد نشود را بیابید.

۵۲) احتمال اینکه در شهر تهران یک نفر در یک روز معین در اثر تصادف با اتومبیل فوت کند ۰/۰۰۰۳ است. یک نمونه $n = ۱۰۰۰۰$ نفری از جمعیت شهر تهران به‌تصادف انتخاب شده است. احتمال اینکه از بین این تعداد در یک روز معین لااقل یک نفر در اثر تصادف فوت کند چقدر است؟

۵۳) اگر Y دارای توزیع یکنواخت در فاصلهٔ $(۰,۵)$ باشد احتمال اینکه

$$۴x^۲ + ۴xY + Y + ۲ = ۰$$

دارای ریشه‌های حقیقی باشد را بیابید.

۵۴) زمان رسیدن اتوبوسی به ایستگاه یک متغیر تصادفی یکنواخت با میانگین ۲ بعد از ظهر و واریانس ۲۱ دقیقه است. حدود زمان رسیدن اتوبوس را به‌دست آورید.

۵۵) در یک نانوایی زمان سرویس به هر مشتری توزیع نمایی با میانگین ۵ دقیقه است. اگر اکنون وارد نانوایی شوید و ۳ نفر در صف باشند، احتمال اینکه کمتر از ۲۰ دقیقه نوبت شما شود چقدر است؟

۵۶) در بیمارستانی نوزادها با نرخ پواسون ۱۲ نفر در روز متولد می‌شوند. احتمال اینکه حداقل ۷ ساعت طول بکشد تا ۳ نوزاد متولد شوند چقدر است؟

۵۷) فرض کنید X زمان تعمیر یک تلویزیون برحسب ساعت باشد که توزیع نمایی با میانگین ۲ ساعت دارد.

(الف) اگر در یک روز یک تعمیرکار تلویزیون‌ها را یک‌به‌یک و مستقلاً تعمیر کند. احتمال اینکه سومین تلویزیون تعمیری اولین تلویزیونی باشد که قبل از ۲ ساعت تعمیر می‌شود را بیابید.

(ب) اگر بدانیم که در ۴ ساعت حداقل ۲ تلویزیون تعمیر شده است. احتمال اینکه در همان ۴ ساعت حداکثر ۳ تلویزیون تعمیر شده باشد را بیابید.

۵۸) اگر $X \sim G(p)$ و $Y \sim E(\theta)$ آنگاه θ را بر حسب p چنان تعیین کنید که $P(X > ۱) = P(Y > ۱)$.

۵۹) خریداری یک محموله بزرگ از کالا را به شرطی می‌پذیرد که در یک نمونه ۱۰۰۰ تایی از این محموله، حداکثر ۱۰ کالای معیوب وجود داشته باشد. اگر در تمام محموله ۰/۵ درصد کالای معیوب وجود داشته باشد.

احتمال اینکه محموله پذیرفته نشود چقدر است؟

(الف) با به‌کار بردن توزیع پواسون،

(ب) با به‌کار بردن منحنی نرمال به عنوان تقریبی از توزیع دوجمله‌ای

۶۰) معمولاً در هر ساعت ۱۲۰ بار به آژانسی تلفن می‌شود. احتمال‌های زیر را محاسبه کنید.

(الف) در یک دقیقه آینده حداقل ۳ تماس تلفنی به آژانس مد نظر برقرار شود.

(ب) فاصلهٔ زمانی بین دو تماس متوالی حداکثر یک دقیقه باشد.

(پ) فاصلهٔ زمانی بین تماس اول و دهم حداقل ۸/۲۳ دقیقه باشد.

۶۱) یک شرکت نوعی چای را در جعبه‌هایی بسته‌بندی می‌کند که وزن خالص چای در

این جعبه‌ها دارای توزیع نرمال با میانگین ۱۰۰ و انحراف معیار ۸ گرم است و وزن جعبه‌ها دارای توزیع نرمال با میانگین ۲۰ و انحراف معیار ۶ گرم است.
(الف) احتمال اینکه وزن خالص چای در یک جعبه از ۹۲ گرم بیشتر باشد را بیابید.
(ب) چند درصد از جعبه‌های چای بسته‌بندی شده در این شرکت وزن کل (وزن جعبه و چای) کمتر از ۱۳۵ گرم دارند؟
(پ) اگر ۴ جعبه را با جای‌گذاری از این جعبه‌های چای بسته‌بندی شده انتخاب کنیم احتمال اینکه حداقل یکی از این جعبه‌ها دارای وزن کل بین ۱۱۰ و ۱۳۰ گرم باشد را بیابید.

۶۲) فرض کنید در یک کارخانه برای تولید و عرضهٔ یک محصول ۳ مرحله ساخت، مونتاژ و بسته‌بندی انجام گیرد که زمان انجام این سه مرحله از یکدیگر مستقل بوده و هر یک دارای توزیع نرمال با میانگین‌های ۳۰، ۱۲ و ۸ دقیقه و انحراف معیارهای ۶، ۳ و ۲ دقیقه به‌ترتیب برای ساخت، مونتاژ و بسته‌بندی باشند.
(الف) اگر ۵ محصول به‌تصادف و با جای‌گذاری از محصولات این کارخانه انتخاب شوند. احتمال اینکه حداقل ۲ عدد از این محصولات دارای زمان تولید و عرضه (ساخت، مونتاژ و بسته‌بندی) بین ۴۳ و ۵۷ دقیقه باشند را بیابید.
(ب) اگر ۵۰ محصول به‌تصادف و با جای‌گذاری از محصولات این کارخانه انتخاب شوند، احتمال اینکه حداکثر ۲۰ عدد از این محصولات دارای زمان تولید و عرضه بیش از ۵۰ دقیقه باشند را بیابید.

۶۳) در یک کلاس ۷۰ نفری نمرات امتحان درس آمار دارای توزیع نرمال با میانگین نامعلوم μ و انحراف معیار ۱۰ است.
(الف) اگر ۳۳% از افراد این کلاس نمره‌اشان کمتر از ۴۵/۶ باشد، μ را پیدا کنید.
(ب) چند نفر در این کلاس نمره‌اشان بیش از ۷۵ می‌شود.
(پ) اگر علاوه بر این امتحان، درس آمار پروژه‌ای داشته باشد که نمرهٔ آن دارای توزیع نرمال با میانگین ۱۵ و انحراف معیار ۲ است و کسی در درس آمار قبول می‌شود که جمع نمره‌های امتحان و پروژه‌اش بیشتر از ۷۰ باشد، چند درصد در این کلاس قبول می‌شوند؟

۶۴) طول عمر یک قطعهٔ بخصوص برحسب ساعت دارای تابع چگالی زیر است.

$$f(t) = \begin{cases} ۲e^{-۲t} & t > ۰ \\ ۰ & \text{سایر نقاط} \end{cases}$$

(الف) میانگین عمر این قطعه را به‌دست آورید.

(ب) اگر پنج قطعه از این قطعه‌ها را در پنج دستگاه مختلف که به‌طور مستقل کار می‌کنند، نصب کنیم احتمال اینکه دقیقاً ۲ قطعه از این ۵ قطعه قبل از ۸ ساعت از بین بروند را بیابید.

۶۵) اگر X و Y متغیرهای مستقل هندسی با میانگین $\frac{۱}{p}$ باشند، $P(X = Y)$ را بیابید.

۶۶) فرض کنید X و Y دو متغیر تصادفی مستقل از یکدیگر با توزیع یکسان یکنواخت گسسته روی مجموعهٔ $\{۱, ۲, \cdots, \theta\}$ هستند. مقدار $P(X \neq Y)$ را بیابید.

۶۷) فرض کنید X و Y دو متغیر تصادفی با توزیع‌های نمایی با میانگین‌های به‌ترتیب $\lambda_۱^{-۱}$ و $\lambda_۲^{-۱}$ هستند. اگر $T = \min(X, Y)$. مقدار $P(T = X)$ را بیابید.

۶۸) متغیر تصادفی X دارای توزیع وایبول است، اگر و تنها اگر تابع چگالی احتمال آن به‌صورت

$$f(x) = \begin{cases} kx^{\beta-۱}e^{-\alpha x^{\beta}} & x > ۰ \\ ۰ & x < ۰ \end{cases}$$

باشد، که در آن $\alpha > ۰$ و $\beta > ۰$.

(الف) k را بر حسب α و β بیان کنید.

(ب) نشان دهید که $\mu = \alpha^{-\frac{۱}{\beta}}\Gamma\left(۱ + \frac{۱}{\beta}\right)$

۶۹) اگر X توزیع یکنواخت در فاصله $(-۱, ۳)$ داشته باشد و Y دارای توزیع‌نمایی با پارامتر θ باشد، مقدار θ را چنان تعیین کنید که $\sigma_X^۲ = \sigma_Y^۲$.

۷۰) متغیر تصادفی X دارای توزیع پارتو است، اگر و تنها اگر چگالی آن به‌صورت

$$f(x) = \begin{cases} \dfrac{\alpha}{x^{\alpha+۱}} & x > ۱ \\ ۰ & x < ۱ \end{cases}$$

باشد، که در آن $\alpha > ۰$.

(الف) نشان دهید که μ_r تنها وقتی وجود دارد که $r < \alpha$.

(ب) نشان دهید که برای توزیع پارتو، $\mu = \frac{\alpha}{\alpha - ۱}$ به شرط آنکه $\alpha > ۱$ باشد.