توزیع های گسسته خاص:

توزیع برنولی:

آزمایش برنولی: اگر فضای نمونه یک آزمایش تصادفی فقط دو برآمد موفقیت و شکست باشد، آزمایش را یک آزمایش برنولی گوییم.

به عنوان مثال نتیجه پرتاب یک سکه، نتیجه آزمایش های پزشکی که جواب آن به صورت منفی و مثبت است، پرتاب یک تاس وقتی ظاهر شدن عدد شش موفقیت و بقیه شکست هستند، ... همگی نمونه هایی از آزمایش برنولی هستند.

تعریف: اگر X تعداد موفقیت ها در یک آزمایش برنولی باشد، در این صورت X را یک متغیر تصادفی برنولی می‌نامیم. اگر p احتمال پیروزی و $q=1-p$ احتمال شکست در یک آزمایش برنولی باشد، در این صورت تابع جرم احتمال متغیر برنولی X به صورت زیر است:

$$f_X(x)=\begin{cases} p^x q^{1-x} & x=0,1 \\ 0 & ow \end{cases}$$

قضیه: اگر X یک متغیر تصادفی برنولی باشد،

$$E[X]=p, \quad \sigma_X^2=pq$$

$$\begin{aligned} E[X] &= 0\times q+1\times p=p \\ E[X^2] &= 0\times q+1^2\times p=p \\ \sigma_X^2 &= E[X^2]-E[X]^2=p-p^2=p(1-p)=pq \end{aligned}$$

توزیع دوجمله ای:

فرض کنید n آزمایش تصادفی برنولی را به طور مستقل از هم انجام داده ایم. اگر X را تعداد موفقیت ها در این n آزمایش برنولی تعریف کنیم، X را متغیر تصادفی دوجمله‌ای می‌نامیم. تابع جرم احتمال این توزیع به صورت زیر است:

$$f_X(x)=\binom{n}{x}p^x q^{n-x} \quad x=0,1,2,\dots,n$$

قضیه: اگر X یک متغیر تصادفی دوجمله‌ای باشد،

$$E[X]=np, Var[X]=npq$$

اثبات: می توان X را به صورت

$$X=Y_1+Y_2+\dots+Y_n$$

در نظر گرفت که در آن Y_iها متغیرهای تصادفی برنولی مربوط به آزمایش برنولی iام هستند. لذا:

$$E[X]=\sum_{i=1}^{n}E[Y_i]=np$$

از طرفی با توجه به اینکه Y_iها از یکدیگر مستقل هستند، داریم:

$$Var[X]=\sum_{i=1}^{n}Var[Y_i]=npq$$

مثال: یک چراغ کنترل ترافیک چنان است که ۴۰ ثانیه قرمز، ۲۰ ثانیه سبز، و ۱۰ ثانیه زرد است. رانندهای از این چراغ ترافیک ۸ بار عبور می‌کند. احتمال اینکه ۳ بار با چراغ قرمز مواجه شود چقدر است؟ احتمال اینکه کمتر از ۵ بار به چراغ قرمز برخورد کند چقدر است؟

حل: فرض کنید X تعداد دفعاتی باشد که فرد مورد نظر به چراغ راهنمایی برخورد می‌کند.

$$p=\frac{40}{10+20+40}=\frac{4}{7}$$
$$q=1-p=\frac{3}{7}$$
$$P(X=3)=\binom{7}{3}\left(\tfrac{4}{7}\right)^3\left(\tfrac{3}{7}\right)^4$$
$$P(X<5)=\sum_{i=0}^{4}\binom{7}{i}\left(\tfrac{4}{7}\right)^i\left(\tfrac{3}{7}\right)^{7-i}$$

مثال: فردی به طور کاملا تصادفی در حال قدم زدن است. وی با احتمال $\frac{1}{3}$ به سمت راست و با احتمال $\frac{2}{3}$ به سمت چپ قدم می‌زند. الف) اگر این فرد صد قدم بردارد، احتمال اینکه در نهایت سر جای اولش باشد چقدر است؟ ب) احتمال اینکه ۲۰ قدم به سمت راست حرکت کرده باشد چقدر است؟

حل:

الف) X: تعداد قدم هایی که به سمت راست گام بر می‌دارد.

Y: تعداد قدم هایی که به سمت چپ بر می‌دارد.

$$\begin{cases}x+y=100\\ x-y=0\end{cases}$$

با حل این معادله $x=50$ به دست می‌آید. حال کافیست $P(X=50)$ را محاسبه کنیم:

$$P\left(X=50\right)=\binom{100}{50}\left(\tfrac{1}{3}\right)^{50}\left(\tfrac{2}{3}\right)^{50}$$

حل ب)

$$\begin{cases} x+y=100 \\ x-y=20 \end{cases}$$

که جواب این دستگاه $x=60$ است. پس جواب مسئله به صورت زیر است:

$$P\left(X=60\right)=\binom{100}{50}\left(\tfrac{1}{3}\right)^{60}\left(\tfrac{2}{3}\right)^{40}$$

توزیع هندسی:

فرض کنید یک آزمایش برنولی را اینقدر انجام دهیم تا برای اولین بار پیروزی ظاهر شود و X را تعداد آزمایش ها تا کسب اولین موفقیت تعریف کنیم. در این صورت متغیر تصادفی X را یک متغیر تصادفی هندسی می‌نامیم. تابع جرم احتمال این توزیع به صورت زیر است:

$$f_X(x)=pq^{x-1} \quad x=1,2,\dots$$

قضیه: امید ریاضی و واریانس متغیر تصادفی هندسی به صورت زیر است:

$$E[X]=\frac{1}{p} \quad \sigma^2=\frac{q}{p^2}$$

مثال: در مثال قدم زدن تصادفی احتمال اینکه فرد مورد نظر برای اولین بار در پنجمین قدم خود به سمت راست قدم بردارد چقدر است؟

$$p=\frac{1}{3},q=1-p=\frac{2}{3}$$

$$P\left(X=5\right)=pq^4=\frac{1}{3}\left(\frac{2}{3}\right)^4=\frac{16}{243}$$

توزیع دوجمله ای منفی:

فرض کنید یک آزمایش برنولی را اینقدر ادامه دهیم تا k امین پیروزی رخ دهد. اگر X تعداد آزمایش های لازم جهت رخ دادن k امین پیروزی تعریف گردد، X را متغیر تصادفی دوجمله‌ای منفی می‌نامیم. تابع جرم احتمال متغیر تصادفی دوجمله‌ای منفی به صورت زیر است:

$$f_X\left(x\right)=P\left(X=x\right)=\binom{x-1}{k-1}p^k q^{x-k} \quad x=k,k+1,\dots$$

قضیه: اگر X یک متغیر تصادفی دوجمله ای منفی باشد

$$E[X]=\frac{k}{p} \qquad Var[X]=\frac{k}{p}\left(\frac{1}{p}-1\right)$$

مثال: در مثال کنترل ترافیک، احتمال اینکه راننده در ۷ امین عبور خود برای سومین بار به چراغ قرمز برخورد بکند، چقدر است؟

$$k=3$$
$$p=\frac{4}{7}$$
$$P(X=7)=\binom{7-1}{3-1}\left(\frac{4}{7}\right)^3\left(\frac{3}{7}\right)^{7-3}=\binom{6}{2}\left(\frac{4}{7}\right)^3\left(\frac{3}{7}\right)^4$$

تذکر: توزیع هندسی حالت خاصی از توزیع دوجمله ای منفی است که درآن $k=1$.

توزیع فوق هندسی:

یک مجموعه N عضوی را در نظر بگیرید که در آن k شیئ موفقیت و $N-k$ شیئ شکست هستند. حال اگر از این مجموعه یک نمونه n عضوی (بدون جایگذاری انتخاب کنیم) که در آن $n\le N$ و $k\le N$ ، متغیر تصادفی X که تعداد موفقیت ها در این نمونه n تایی است را یک متغیر تصادفی فوق هندسی می‌نامیم و داریم:

$$P(X=x)=\frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \qquad x=\max\{0,n+k-N\},\dots,\min\{k,n\}$$

قضیه: اگر X یک متغیر تصادفی فوق هندسی باشد، مشروط بر اینکه $n=\min\{k,n\}$ ، $0=\max\{0,n+k-N\},$ داریم:

$$E[X]=np,\ \sigma^2{}_X=np(1-p)\frac{N-n}{N-1}$$

که در آن $p=\frac{k}{N}$ ، به $\frac{N-n}{N-1}$ فاکتور تصحیح جمعیت متناهی می‌گویند.

تذکر: با توجه به روابط فوق می‌توان نتیجه گرفت که وقتی جامعه نامتناهی است و یا $n<<N$ در این صورت $\sigma^2=npq$ که همان واریانس توزیع دوجمله ای است. در واقع توزیع دوجمله ای حالت حدی توزیع فوق هندسی است که در آن $p=\frac{k}{n}$ و $n<<N$. (n بیش از $0.05N$ نباشد.)

مثال: احتمال آنکه یک ممیز مالیاتی از بین ۵ اظهار نامه مالیاتی دو اظهارنامه با بخشودگی غیر مجاز بیابد چقدر است به شرط اینکه این ۵ اظهارنامه به تصادف از بین ۱۵ اظهارنامه که شامل ۹ اظهارنامه با بخشودگی غیرمجاز است ، انتخاب شود.

$$P(X=2)=\frac{\binom{9}{2}\binom{6}{3}}{\binom{15}{5}}=0.3853$$

مثال: محموله ای مرکب از ۸۰ دستگاه دزدگیر شامل ۴ دستگاه معیوب است. اگر ۳ دستگاه به تصادف انتخاب شود و برای یک مشتری ارسال گردد، احتمال اینکه مشتری یک دستگاه معیوب دریافت کند چقدر است؟

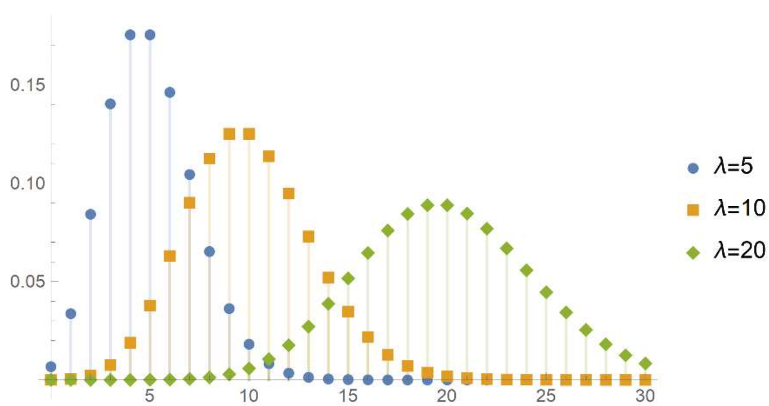
$$p=\frac{4}{80}=0.05$$
$$P(X=1)=\binom{3}{1}(0.05)(0.95)^2=0.1353$$

توزیع پوآسون:

تعریف: گوییم متغیر تصادفی X دارای توزیع پوآسون است هرگاه تابع جرم احتمال آن به صورت زیر باشد:

$$f_X(x)=P(X=x)=\frac{e^{-\lambda}\lambda^x}{x!}\quad x=0,1,2,....$$

که در اینجا λ پارامتر توزیع نام دارد.



شکل ۱ توزیع پوآسون با □ های مختلف

نکته: به سادگی می‌توان نشان داد که تابع فوق یک تابع جرم احتمال است:

$$e^{x} = \sum_{i=0}^{\infty} \frac{x^{i}}{i\,!}$$

$$\sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^{x}}{x\,!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^{x}}{x\,!} = e^{-\lambda} e^{\lambda} = 1$$

قضیه: اگر X یک متغیر تصادفی پوآسون باشد،

$$E[X] = \lambda$$
$$\sigma^2_X = \lambda$$
$$M_X(t) = e^{\lambda(e^t - 1)}$$

$$E[X] = \sum_{x=0}^{\infty} \frac{xe^{-\lambda}\lambda^{x}}{x\,!} = \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^{x}}{(x-1)!} = \lambda \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} = \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

$$E[X(X-1)] = \sum_{x=0}^{\infty} \frac{x(x-1)e^{-\lambda}\lambda^{x}}{x\,!} = \sum_{x=2}^{\infty} \frac{e^{-\lambda}\lambda^{x}}{(x-2)!} = \lambda^2 \sum_{x=2}^{\infty} \frac{e^{-\lambda}\lambda^{x-2}}{(x-2)!} = \lambda^2 e^{-\lambda} e^{\lambda} = \lambda^2$$

$$E[X(X-1)] = E[X^2 - X] = E[X^2] - E[X]$$
$$E[X^2] = \lambda^2 + \lambda$$
$$\sigma^2 = E[X^2] - E[X]^2 = \lambda$$

$$M_X(t) = E[e^{tX}] = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^{x}}{x\,!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\left(e^t \lambda\right)^x}{x\,!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda\left(e^t - 1\right)}$$

کاربرد: فرض کنید X تعداد رویدادهای خاصی باشد که در طول بازه (زمانی یا مکانی) مشخصی مانند T اتفاق می‌افتد و به طور متوسط تعداد این رویدادها در این بازه برابر λ باشد. در این صورت اگر شرایط زیر برقرار باشد، متغیر تصادفی X یک متغیر تصادفی پوآسون است:

۱) رویداد ها به طور مستقل از هم رخ دهند و اطلاع از یک رویداد هیچ اطلاعی از زمان یا مکان رویداد بعدی ندهد.

۲) در فواصل به اندازه کافی کوتاه Δt ، احتمال رخ دادن یک رویداد فقط به طول Δt بستگی داشته باشد.

۳) در فواصل به اندازه کافی کوتاه Δt ، امکان رخ دادن دو رویداد ناچیز باشد.

در زیر به بیان مثال هایی می‌پردازیم.

- ✓ تعداد بازدیدکنندگان یک سایت در یک ساعت
- ✓ تعداد تماس های تلفنی به یک منشی در یک روز

- ✓ تعداد تولدها، ازدواج‌ها، و ... در طول یک بازه زمانی
- ✓ تعداد تصادفات رانندگی در طول یک روز در یک خیابان
- ✓ تعداد زدگی های یک پارچه در ۵ متر از آن
- ✓ تعداد بیماران مراجعه کننده به یک کلینیک

مثال: به یک کلینیک به طور متوسط در هر ساعت ۱۰ بیمار مراجعه می‌کنند. احتمال اینکه:

الف) در یک ساعت ۳ بیمار

ب) بیشتر از ۴ بیمار

ج) در نیم ساعت ۳ بیمار وارد شوند را بیابید.

الف)

$$\begin{aligned} &T = 1 \\ &\lambda = 10 \\ &P(X = 3) = \frac{e^{-10} 10^3}{3!} = 0.0075 \end{aligned}$$

ب)

$$P(X \geq 5) = 1 - P\left(X \leq 4\right) = 1 - \sum_{k=0}^{4} \frac{e^{-10} 10^k}{k\,!} = 0.97$$

ج)

$$\begin{aligned} &T = 1 \Rightarrow \lambda = 10 \\ &T = .5 \Rightarrow \lambda = 5 \\ &P(x = 3) = \frac{e^{-5} 5^3}{3!} = 0.14 \end{aligned}$$

مثال: پارچه تولیدی در یک کارخانه نساجی به طور متوسط در هر ۲۰ متر ۵ زدگی دارد. یک پارچه به طول ۱۰ متر از این کارخانه خریداری می‌کنیم. احتمال اینکه این پارچه ۳ زدگی داشته باشد را محاسبه کنید.

حل:

$$\begin{aligned} &T = 20 \ \lambda = 5 \\ &T = 10 \ \ \lambda = 2.5 \\ &\lambda = 2.5 \\ &P(X = 3) = \frac{e^{-2.5} (2.5)^3}{3!} = 0.213 \end{aligned}$$

تقریب توزیع دوجمله ای با توزیع پوآسون: اگر p به اندازه کافی کوچک و n بزرگ باشد می‌توان توزیع دوجمله ای را به کمک توزیع پوآسون با پارامتر $\lambda = np$ تخمین زد.

به طور خلاصه باید شرایط زیر جهت این تقریب برقرار باشد:

۱) $np < 10$
۲) $p < 0.05$
۳) $n \geq 20$ و اگر $n \geq 100$ عالی

مثال: در شهری معین ۰/۰۴ رانندگان در هر سال درگیر تصادفات اتوموبیل می‌شوند. احتمال اینکه از بین ۱۵۰ راننده :

الف) تنها ۵ نفر در سال

ب) حداکثر ۳ نفر تصادف داشته باشند چقدر است؟

حل:

الف) چون $n \geq 100$ و $\lambda = np = 6 < 10$ همچنین $p = 0.04 < 0.05$ لذا از توزیع پوآسون استفاده می‌کنیم.

$$P(X = 5) = \frac{e^{-6} 6^5}{5!} = 0.1606$$

$$P\left(X \leq 3\right) = \sum_{k=0}^{3} \frac{e^{-6} 6^k}{k!} = 0.16$$

توزیع چند جمله ای :

تعریف: فرض کنید یک آزمایش تصادفی فقط شامل k برآمد مانند $e_1, e_2, \ldots, e_k$ و احتمال رخ دادن هر یک برابر با به ترتیب $p_1, p_2, \ldots, p_k$ باشد. حال اگر این آزمایش را n بار مستقل از هم انجام دهیم و متغیر های تصادفی X_i، که $1 \leq i \leq k$ ، تعداد رخداد برآمد e_i در این n آزمایش باشد. این متغیر های تصادفی دارای چگالی احتمال توام چندجمله‌ای به صورت زیر هستند:

$$P\left(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k\right) = \frac{n!}{x_1! x_2! \ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$$
$$x_1 + x_2 + \ldots + x_k = n$$

مثال: جدول زیر فراوانی نسبی گروه های خونی را نشان می‌دهد:

AB	B	A	O
0.04	0.10	0.42	0.44

احتمال اینکه در یک گروه ۱۰ نفره ۳ نفر گروه خونی A، ۴ نفر گروه خونی O، یک نفر B، و دو نفر AB باشند را به دست آورید.

$$P\left(X_1=4, X_2=3, X_3=1, X_4=2\right)=\frac{10!}{4!3!2!}(0.44)^4(0.42)^3(0.1)(0.04)^2$$

تهیه و تنظیم: امیرحسین سبحانی ـ دانشکده ریاضی دانشگاه سمنان