آمار و احتمال مهندسی

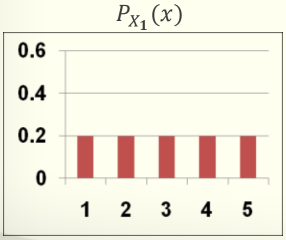
واریانس و توزیع‌های گسسته مهم

(Ross 4.5-4.7)

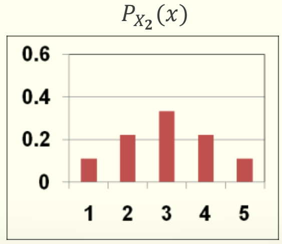
واریانس (Variance)

- میانگین نشان می‌داد که مقادیر X حول و حوش چه مقداری هستند.
- کمیت دیگری لازم داریم که میزان پراکندگی مقادیر X حول مقدار میانگین را به ما نشان دهد.
- واریانس بیانگر میزان پراکندگی احتمال حول مقدار متوسط (میانگین) است.
- میانگین‌ها یکسان ($E(X_i)=3$)، ولی پراکندگی‌ها متفاوت:

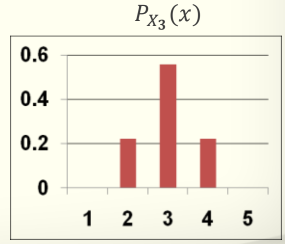
$P_{X_1}(x)$



$P_{X_2}(x)$



$P_{X_3}(x)$



واریانس

- فاصله متغیر تصادفی X از میانگین آن برابر است با: $|X - \mathrm{E}(X)|$
- طبق تعریف توزیع یا واریانس متغیر تصادفی X به صورت زیر است:

$$\mathrm{var}(X) = \mathrm{E}\left(\left(X - \mathrm{E}(X)\right)^2\right)$$

- کار با تعریف بالا ساده‌تر از کار با $\mathrm{E}(|X - \mathrm{E}(X)|)$ است.
- واریانس کمیتی نامنفی است: $\mathrm{var}(X) \geq 0$
- لذا متداول است که آن را با σ_X^2 یا σ^2 نشان می‌دهند:

$$\sigma^2 = E\left((X - \mu)^2\right)$$

گشتاور مرکزی مرتبه دوم

انحراف معیار

- σ را انحراف معیار (Standard Deviation) می‌نامند.
- توجه کنید که $\mathrm{var}(X)$ از جنس مربع متغیر تصادفی است، ولی σ از جنس خود متغیر تصادفی است.
- از تعریف امید ریاضی برای متغیر تصادفی گسسته نتیجه می‌شود که:

$$\sigma^2 = \sum_i (x_i - \mu)^2 P(x_i)$$

خواص واریانس

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

اثبات:

$$\sigma^2 = \mathrm{E}\left((X-\mu)^2\right) = \mathrm{E}\left(X^2 - 2\mu X + \mu^2\right)$$

$$\sigma^2 = \mathrm{E}\left(X^2\right) - 2\mu\mathrm{E}(X) + \mu^2 = \mathrm{E}\left(X^2\right) - \mu^2$$

- نتیجه فرعی:

$$\mathrm{E}\left(X^2\right) = \sigma^2 + \mu^2 \Rightarrow \mathrm{E}\left(X^2\right) \geq \mathrm{E}^2(X)$$

- $\mathrm{E}(X^2)$ را میانگین مربعی (mean square) گویند و داریم:

$$\sqrt{E(X^2)} = \text{root mean square (rms)}$$

مثال

- دیدیم که اگر X متغیر تصادفی خروجی پرتاب یک تاس باشد داریم: $E(X) = 7/2$
- از سوی دیگر:

$$E(X^2) = (1^2)\times\frac{1}{6} + (2^2)\times\frac{1}{6} + (3^2)\times\frac{1}{6} + (4^2)\times\frac{1}{6} + (5^2)\times\frac{1}{6} + (6^2)\times\frac{1}{6}$$
$$= \frac{91}{6}$$

$$\mathrm{var}(X) = E\left(X^2\right) - E(X)^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$$

خواص واریانس

۲) اگر $Y = aX + b$ باشد، داریم: $\mathrm{var}(Y) = a^2\mathrm{var}(X)$

اثبات:

$$\begin{aligned}\mathrm{var}(Y) &= E\big((Y - E[Y])^2\big)\\ &= E\big((aX + b - aE[X] - b)^2\big)\\ &= E\big(a^2(X - E[X])^2\big)\\ &= a^2E\big((X - E[X])^2\big) = a^2\mathrm{var}(X)\end{aligned}$$

توزیع برنولی

Jacob Bernoulli
(1654-1705)

- اگر خروجی یک آزمایش تصادفی فقط دو حالت موفقیت و شکست را داشته باشد، می‌توانیم این خروجی را با یک متغیر تصادفی شاخص X مدل کنیم:

  شکست : 0 و موفقیت : 1

  $P(X = 0) = 1 - p$ و $P(X = 1) = p$

- X را یک متغیر تصادفی برنولی می‌نامیم: $X \sim \mathrm{Ber}(p)$

$$E[X] = p \;,\; \mathrm{Var}(X) = p(1 - p)$$

- مثال: پرتاب سکه، مذکر یا مونث بودن نوزاد، پاس شدن درس آمار و احتمال، دوست داشتن یک فیلم روی Netflix

توزیع دوجمله‌ای

- اگر یک آزمایش تصادفی با دو خروجی موفقیت (با احتمال p) و شکست (با احتمال $1-p$) را n بار تکرار کنیم، و متغیر تصادفی X را به این صورت تعریف کنیم که:

$X =$ تعداد موفقیت‌ها در n آزمایش

آنگاه X دارای توزیع دوجمله‌ای (binomial) خواهد بود:

$X \sim \mathrm{Bin}(n, p)$

$$P_X(k) = P\{X = k\} = \binom{n}{k} p^k q^{n-k} \quad : \quad k = 0, 1, 2, \ldots, n$$

و برای سایر مقادیر k داریم: $P(k) = 0$

- توزیع دوجمله‌ای در حالت خاص $n = 1$ تبدیل به توزیع برنولی می‌شود.

رابطه توزیع برنولی و توزیع دوجمله‌ای

- متغیر تصادفی دوجمله‌ای را می‌توانیم به صورت مجموع n متغیر تصادفی مستقل برنولی در نظر بگیریم:

$$X = X_1 + X_2 + \cdots + X_n$$

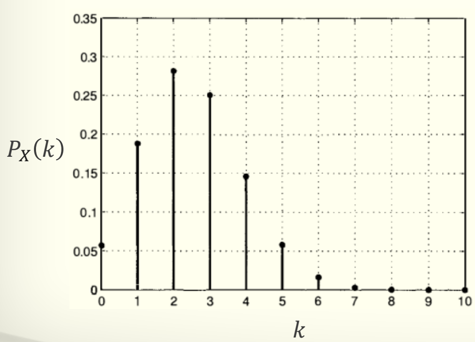
$$X_i \sim \mathrm{Ber}(p) \ , X \sim \mathrm{Bin}(n, p)$$

- مثال:
  - تعداد شیرها در پرتاب n سکه
  - تعداد 1 ها در یک رشته n بیتی تصادفی
  - تعداد دیسک‌های خراب‌شده در یک کلاستر ۱۰۰۰ کامپیوتری
  - تعداد آرای یک کاندیدا در انتخابات

ویژگی‌های توزیع دوجمله‌ای

- ویژگی‌های یک متغیر تصادفی با توزیع دوجمله‌ای:
  - تعداد آزمایش‌های تصادفی ثابت است: n
  - آزمایش‌های تصادفی مستقل از هم هستند.
  - هر آزمایش دو خروجی ممکن دارد: شکست و موفقیت
  - احتمال موفقیت در همه آزمایش‌ها یکسان است: p

$X \sim \text{Bin}(10, 0.25)$:

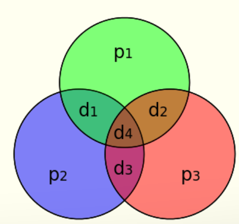


مثال: کد تصحیح خطا

- کدهای تصحیح خطا برای تشخیص و تصحیح خطاهای رخ‌داده در دنباله بیتی عبوری از یک کانال مخابراتی نویزی به کار می‌روند.
- یکی از رایج‌ترین کدهای تصحیح خطا، کد Hamming(7,4) است. در این کد پس از هر چهار بیت داده $d_1d_2d_3d_4$، سه بیت توازن $p_1p_2p_3$ اضافه می‌شود:

$$p_1 = d_1 \oplus d_2 \oplus d_4$$
$$p_2 = d_1 \oplus d_3 \oplus d_4$$
$$p_3 = d_2 \oplus d_3 \oplus d_4$$

$1100 \rightarrow 1100011$



مثال: کد تصحیح خطا

- اگر هر بیت از اطلاعات با احتمال 0.1 خراب شود، احتمال دریافت صحیح یک دنباله چهار بیتی چقدر است؟
- متغیر تصادفی X را تعداد بیت‌های خراب شده در دنباله ۴-بیتی تعریف می‌کنیم.
- واضح است که X دارای توزیع دوجمله‌ای است: $X \sim \text{Bin}(4,0.1)$
- دریافت صحیح دنباله به معنای پیشامد $\{X = 0\}$ است:

$$P\{X = 0\} = \binom{4}{0}(0.1)^0(1 - 0.1)^{4-0} = 0.6561$$

مثال: کد تصحیح خطا

- اگر هر بیت از اطلاعات با احتمال 0.1 خراب شود، احتمال دریافت صحیح یک دنباله هفت بیتی که با استفاده از کد Hamming(7,4) ساخته شده، چقدر است؟
- متغیر تصادفی X را تعداد بیت‌های خراب شده در دنباله ۷-بیتی تعریف می‌کنیم.
- X دارای توزیع دوجمله‌ای است: $X \sim \text{Bin}(7,0.1)$
- دیدیم که اگر یک بیت خطا در این دنباله اتفاق بیافتد، این خطا قابل تشخیص و تصحیح است، بنابراین دریافت صحیح دنباله به معنای پیشامد $\{X = 0\} \cup \{X = 1\}$ است:

$$P\{X = 0 \cup X = 1\} = \binom{7}{0}(0.1)^0(1 - 0.1)^{7-0} + \binom{7}{1}(0.1)^1(1 - 0.1)^{7-1}$$
$$\approx 0.8503$$

میانگین توزیع دوجمله‌ای

- راه اول: استفاده از خاصیت خطی بودن امید ریاضی
- دیدیم که متغیر تصادفی دوجمله‌ای $X \sim \text{Bin}(n,p)$ را می‌توان به صورت مجموع n متغیر تصادفی برنولی با احتمال موفقیت p نوشت:

$$X = X_1 + X_2 + \cdots + X_n \quad : \quad X_i \sim \text{Ber}(p)$$

- بنابراین با استفاده از خاصیت خطی بودن امید ریاضی داریم:

$$E[X] = E[X_1] + E[X_2] + \cdots + E[X_n] = np$$

راه دوم: استفاده از اتحاد جابجایی

$$k\binom{n}{k} = n\binom{n-1}{k-1}$$

میانگین توزیع دوجمله‌ای

$$E[X] = \sum_{k=0}^{n} k\binom{n}{k} p^k q^{n-k}$$

$$= \sum_{k=1}^{n} n\binom{n-1}{k-1} p \times p^{k-1} q^{n-k}$$

$$= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} q^{(n-1)-(k-1)} \qquad (k-1) \to m$$

$$= np \sum_{m=0}^{n-1} \binom{n-1}{m} p^m q^{(n-1)-m} = np(p+q)^{n-1} = np \times 1 = np$$

واریانس توزیع دوجمله‌ای

$$E[X^2] = \sum_{k=0}^{n} k^2 \binom{n}{k} p^k q^{n-k} = \sum_{k=1}^{n} kn \binom{n-1}{k-1} p \times p^{k-1} q^{n-k}$$

$$= np \sum_{k=1}^{n} k \binom{n-1}{k-1} p^{k-1} q^{(n-1)-(k-1)} \qquad (k-1) \to m$$

$$= np \sum_{m=0}^{n-1} (m+1) \binom{n-1}{m} p^m q^{(n-1)-m}$$

$$= np \left(\sum_{m=0}^{n-1} m \binom{n-1}{m} p^m q^{(n-1)-m} + \sum_{m=0}^{n-1} \binom{n-1}{m} p^m q^{(n-1)-m} \right)$$

واریانس توزیع دوجمله‌ای

$$E[X^2] = np \left(\sum_{m=0}^{n-1} m \binom{n-1}{m} p^m q^{(n-1)-m} + \sum_{m=0}^{n-1} \binom{n-1}{m} p^m q^{(n-1)-m} \right)$$

$$= np(E[Y \sim Bin(n-1,p)] + (p+q)^{n-1})$$

$$= np\big((n-1)p + 1\big) = (np)^2 + np(1-p)$$

$$Var(X) = E[X^2] - (E[X])^2 = (np)^2 + np(1-p) - (np)^2$$

$$Var(X) = np(1-p) = npq$$

حالت خاص توزیع دوجمله‌ای

- در مثال کد تصحیح خطا، با دنباله ۴-بیتی و احتمال خطای بیتی 0.1 برخورد کردیم.
  - در عمل با رشته‌های بیتی بسیار بزرگتر ($n = 10^4$) و احتمال خطای بسیار کوچکتر ($p = 10^{-6}$) مواجه هستیم.
- محاسبه احتمالات برای متغیر تصادفی $X \sim \mathrm{Bin}(10^4, 10^{-6})$ بسیار دشوار است.
- در عمل در بسیاری از مواقع با این حالت (n خیلی بزرگ و p خیلی کوچک) برخورد می‌کنیم:
  - تعداد بیت‌های خطادار در فایل‌های روی دیسک
  - تعداد بازدیدکنندگان یک وبسایت محبوب
  - تعداد سرورهای از کار افتاده در یک مرکز داده بسیار بزرگ در یک روز

حالت حدی توزیع دوجمله‌ای

- قضیه: اگر p خیلی کوچک، n خیلی بزرگ، و $\lambda = np$ عددی متوسط (حدودا بین ۱ تا ۱۰) باشد، خواهیم داشت:

$$P_n(k) = \binom{n}{k} p^k (1-p)^{n-k} \cong e^{-np} \frac{(np)^k}{k!}$$

- به عبارت دیگر اگر $p \to 0$، و $np \to \lambda$:

$$\lim_{n\to\infty} P_n(k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

اثبات قضیه

اثبات:

$$np = \lambda \Rightarrow p = \frac{\lambda}{n}$$

$$P_n(k) = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{k!\,(n-k)!}\left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$P_n(k) = \frac{\lambda^k}{k!} \times \frac{n!}{(n-k)!\,n^k} \times \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k} = \frac{\lambda^k}{k!} \times \frac{n(n-1)\dots(n-k+1)}{n^k} \times \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k}$$

$$\lim_{n\to\infty} P_n(k) = \frac{\lambda^k}{k!} \times \frac{n^k}{n^k} \times \frac{e^{-\lambda}}{(1-0)^k} = e^{-\lambda}\frac{\lambda^k}{k!}$$

توزیع پواسون

- متغیر تصادفی X را دارای توزیع پواسون با پارامتر λ گوییم اگر:

$$P\{X = k\} = e^{-\lambda}\frac{\lambda^k}{k!} \quad : \quad k = 0, 1, 2, \dots, +\infty$$

و به صورت $X \sim \text{Poi}\ (\lambda)$ نمایش می‌دهیم.

- توجه کنید که از بسط تیلور e^{λ} داریم:

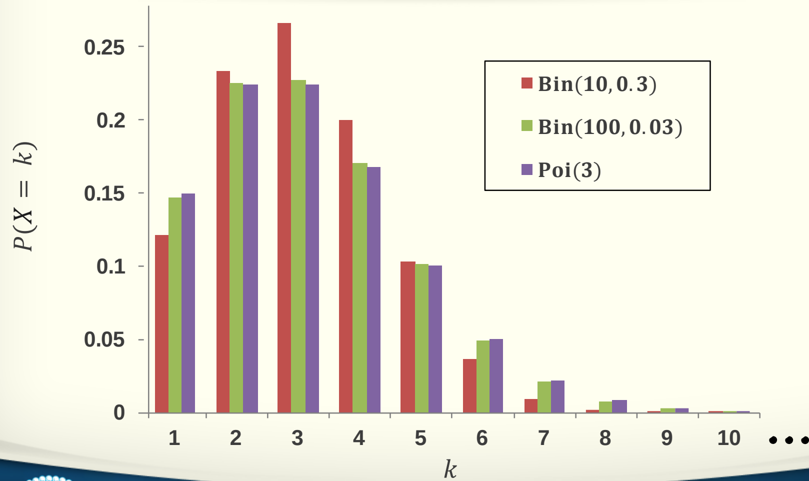
$$e^{\lambda} = \frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!} + \cdots = \sum_{i=0}^{\infty}\frac{\lambda^i}{i!}$$

Siméon Poisson (1781–1840)

بنابراین:

$$\sum_{i=0}^{\infty} P(X = i) = \sum_{i=0}^{\infty} e^{-\lambda}\frac{\lambda^i}{i!} = e^{-\lambda}\sum_{i=0}^{\infty}\frac{\lambda^i}{i!} = e^{-\lambda}e^{\lambda} = 1$$

مقایسه پواسون و دوجمله‌ای



مثال

- برای یک دنباله بیتی به اندازه $n = 10^4$ و احتمال خطای کوچک $p = 10^{-6}$ برای یک بیت، احتمال دریافت صحیح دنباله چقدر است؟
- تعداد بیت‌های خطادار در دنباله یک متغیر تصادفی با توزیع $Y\sim\mathrm{Bin}(10^4,10^{-6})$ است که می‌توانیم آن را با یک متغیر تصادفی پواسون با پارامتر $\lambda = 10^4 \times 10^{-6} = 0.01$ تخمین بزنیم: $X\sim\mathrm{Poi}(0.01)$

$$P\{X = 0\} = e^{-0.01}\frac{0.01^0}{0!} = 0.990049834$$

- مقدار دقیق برابر است با:

$$P\{Y = 0\} = \binom{10^4}{0}\left(10^{-6}\right)^0\left(1 - 10^{-6}\right)^{10^4} \approx 0.990049829$$

میانگین و واریانس توزیع پواسون

- برای متغیر تصادفی دوجمله‌ای $X\sim \text{Bin}(n,p)$ داشتیم:

$$E[X] = np$$
$$Var(X) = np(1-p)$$

- می‌دانیم که متغیر تصادفی پواسون $X\sim \text{Poi}(\lambda)$ معادل یک متغیر تصادفی دوجمله‌ای با $n\to\infty$ و $p\to 0$ است که $\lambda = np$:

$$E[X] = np = \lambda$$
$$Var(X) = np(1-p) = \lambda(1-p) = \lambda(1-0) = \lambda$$

- امید ریاضی و واریانس متغیر تصادفی پواسون یکسان و برابر با پارامتر λ است.

میانگین و واریانس توزیع پواسون

- راه دوم: $X\sim Poi(\lambda)$

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

$$= \lambda e^{-\lambda}\sum_{m=0}^{\infty}\frac{\lambda^m}{m!} = \lambda e^{-\lambda}e^{\lambda} = \lambda$$

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

$$= \lambda e^{-\lambda}\sum_{k=1}^{\infty}(k-1+1)\,\frac{\lambda^{k-1}}{(k-1)!}$$

میانگین و واریانس توزیع پواسون

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

$$= \lambda e^{-\lambda}\left(\sum_{k=2}^{\infty} \frac{\lambda^{k-1}}{(k-2)!} + \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}\right)$$

$$= \lambda e^{-\lambda}\left(\lambda \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}\right)$$

$$= \lambda e^{-\lambda}\left(\lambda e^{\lambda} + e^{\lambda}\right) = \lambda(\lambda+1)$$

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

توزیع پواسون

- متغیر تصادفی X را دارای توزیع پواسون با پارامتر λ گوییم اگر:

$$P\{X=k\} = e^{-\lambda}\frac{\lambda^k}{k!} \quad : \quad k = 0, 1, 2, \dots, +\infty$$

و به صورت $X \sim \text{Poi}\ (\lambda)$ نمایش می‌دهیم.

- میانگین توزیع پواسون:

$$E[X] = \lambda$$

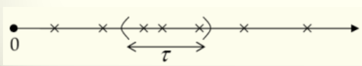
- واریانس توزیع پواسون:

$$Var[X] = \lambda$$

Siméon Poisson
(1781–1840)

کاربرد توزیع پواسون

- تعداد پیشامدهای نادری که در یک بازه زمانی خاص و یا یک محدوده مکانی مشخص اتفاق می‌افتند، غالباً دارای توزیع پواسون هستند:
  - تعداد مکالمات تلفنی در یک مرکز سوئیچ در مدت مشخصی از زمان
  - تعداد نقاط خرابی روی طول مشخصی از یک نوار مغناطیسی یا پارچه
  - تعداد وقوع زلزله یا جنگ در مدت زمانی معین
  - تعداد اشتباهات چاپی در یک صفحه از هر کتاب
  - تعداد خطاهای یک برنامه کامپیوتری با حجم معین که توسط یک فرد نوشته شده است.
  - تعداد ذرات آلفا منتشر شده از یک مادهٔ رادیواکتیو در طی زمان مشخصی
  - تعداد بسته‌های داده دریافتی در یک مسیریاب شبکه اینترنت
  - تعداد بیمارانی که در هر ساعت به بخش اورژانس یک بیمارستان می‌رسند.
- در این آزمایش‌ها $n \to \infty$، $p \to 0$ ولی $np \to \lambda$:



محاسبه سریع احتمالات پواسون

- فرض کنید $X \sim \mathrm{Poi}(\lambda)$:
  - می‌خواهیم احتمال $P\{X = i\}$ را برای چندین مقدار i محاسبه کنیم، برای مثال:

$$F_X(a) = P(X \le a) = \sum_{i=0}^{a} P(X = i)$$

- محاسبه $P\{X = i + 1\}$ از روی $P\{X = i\}$:

$$\frac{P\{X = i + 1\}}{P\{X = i\}} = \frac{e^{-\lambda}\lambda^{i+1}/(i+1)!}{e^{-\lambda}\lambda^{i}/i!} = \frac{\lambda}{i+1}$$

- بنابراین می‌توانیم از روابط بازگشتی زیر استفاده کنیم:

$$P\{X = 0\} = \frac{e^{-\lambda}\lambda^0}{0!} = e^{-\lambda} \quad , \qquad P\{X = i + 1\} = \frac{\lambda}{i+1} P\{X = i\}$$

تقریب پواسون

- دیدیم که برای مدل کردن یک پیشامد با متغیر تصادفی دوجمله‌ای، باید شرایطی از قبیل مستقل بودن آزمایش‌ها و ثابت بودن احتمال موفقیت فراهم باشد.
- نکته جالب توجه درباره تقریب پواسون این است که حتی اگر از برخی از این شرایط به مقدار اندکی تخطی شود، باز هم تقریب دقیقی از احتمال مورد نظر به دست می‌آید.
- در موارد زیر می‌توانیم از متغیر تصادفی پواسون برای محاسبه احتمالات استفاده کنیم:
  - خروجی آزمایش‌ها کاملا مستقل از هم نباشند.
    - مثال: تعداد اشیاء موجود در هر سطر در یک جدول hash بزرگ
  - احتمال موفقیت در هر آزمایش به طور جزئی تغییر کند.
    - تغییرات نسبی کوچک در یک p کوچک
    - مثال: متوسط تعداد درخواست‌هایی که یک سرور وب دریافت می‌کند، ممکن است با توجه به بار شبکه نوسانات جزئی داشته باشد.

مثال: مساله روز تولد

- احتمال این که در بین m نفر، هیچ دو نفری روز تولد یکسان نداشته باشند، چقدر است؟
- $n = \binom{m}{2}$ آزمایش داریم: یک آزمایش برای هر جفت (x,y) که $x \neq y$
- احتمال یکسان بودن تولد (x,y) (احتمال موفقیت) برابر است با:

$$p = 1/365$$

- آزمایش‌ها از هم مستقل نیستند، زیرا یکسان بودن تولد (x,y) مستقل از یکسان بودن تولد (y,z) نیست.
  - بنابراین نمی‌توانیم از توزیع دوجمله‌ای استفاده کنیم.
- از توزیع پواسون برای متغیر تصادفی X (تعداد زوج‌های با روز تولد یکسان) استفاده می‌کنیم:

$$X \sim \text{Poi}(\lambda):\quad \lambda = \binom{m}{2} \times \frac{1}{365} = \frac{m(m-1)}{730}$$

مثال: مساله روز تولد

○ می‌خواهیم احتمال پیشامد $\{X=0\}$ کمتر از 0.5 شود:

$$P\{X=0\}=\frac{e^{-\lambda}\lambda^{0}}{0!}=e^{-\lambda}<0.5$$

$$-\lambda\leq\ln(0.5)\ \Rightarrow\ -\frac{m(m-1)}{730}\leq\ln(0.5)$$

$$m(m-1)\geq-730\ln(0.5)=505.99$$

$$\Rightarrow m\geq 23$$