فصل ۳: آزمایشهای تکراری

Chapter 3

۱. تعبیر مفهومی آزمایشهای تکراری

۲. آزمایش برنولی

۳. قضیهٔ دموآور-لاپلاس

۴. قضیهٔ پواسن

۵. نقاط پواسن

وقتی یک آزمایش تصادفی را تحت شرایط یکسانی تکرار می‌کنیم، دو تعبیر موجود است.

یکی تعبیر تجربی که در آن احتمال واقعهٔ A در فضای Ω حدود $\frac{n_A}{n}$ است.

دیگری تعبیر مفهومی است. در واقع با تکرار آزمایش به جای فضای Ω قبلی یک فضای جدید Ω_n داریم که: $\Omega_n = \underbrace{\Omega \times \cdots \times \Omega}_{\text{بار } n}$.

نقاط این فضای نمونهٔ جدید nتایی‌های مرتبی هستند که هر عنصر آن عضوی از Ω قبلی است.

مثال: پنج بار سکه‌ای را پرتاب می‌کنیم. در این سکه داریم:

$$P\{H\} = p$$

$$P\{T\} = q = 1 - p$$

الف) احتمال واقعهٔ B که در آن فقط دو بار اول شیر بیاید (و بقیه خط) چیست؟

ب) احتمال واقعهٔ D که در آن دو بار شیر بیاید (با هر ترتیبی برای آمدن دو شیر و سه خط) چیست؟

Ω مجموعهٔ کلیهٔ پنج‌تایی‌های مرتبی است که هر عضو این پنج‌تایی‌های مرتب H یا T است:

$$\Omega = \{(HHHHH), \ldots, (TTTTT)\}$$

یعنی $۲^۵$ عنصر دارد که لزوماً متساوی الاحتمال نیستند (چون p لزوماً $\frac{1}{2}$ نیست).

$$B = \{HHTTT\}$$

B = {بار اول شیر بیاید} $\cap$ {بار دوم شیر بیاید} $\cap$ {بار سوم خط بیاید} $\cap$ {بار چهارم خط بیاید} $\cap$ {بار پنجم خط بیاید}

کلیهٔ پنج‌تایی‌های مرتبی که عنصر اول آنها H است

با توجه به فرض استقلال آزمایشها (یعنی احتمال وقایع در یک آزمایش در اثر نتیجهٔ آزمایش دیگر تغییر نمی‌کند) داریم:

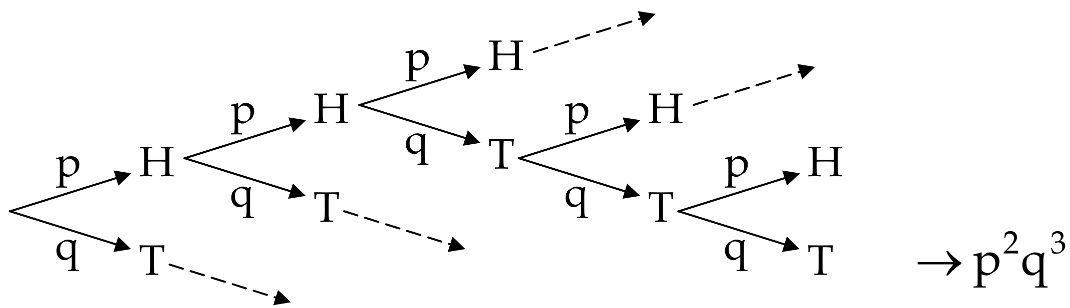
$$P(B) = P(H_1)P(H_2)P(T_3)P(T_4)P(T_5) = p^2q^3$$

اصولاً n آزمایش را مستقل گویند اگر واقعه‌های A_1، A_2، ... و A_n و ... مستقل باشند که A_i واقعه‌ای است که نتیجهٔ آن در ارتباط با آزمایش iام حاصل می‌شود.

اگر بدون توجه به تعریف Ω احتمال این وقایع را در هم ضرب کنیم، بی‌معنی است. چون اشتراکی ندارند.

ضمناً اگر $p = q = \frac{1}{2}$ بود، همان $\frac{1}{2^5}$ می‌شد که قبل از این به دست می‌آوردیم (با تقسیم تعداد نقاط نمونهٔ مطلوب بر کل تعداد نقاط نمونه).

با استفاده از درخت احتمال نیز داریم:



حال برای واقعهٔ D داریم:

$D = \{$دو تا شیر بیاید (و سه خط)$\}$

کلیهٔ وقایع ساده مانند B که دارای دو تا H و سه تا T (با ترتیب مشخص) هستند دارای احتمال p^2q^3 می‌باشند و این وقایع از هم جدا هستند و D در واقع اجتماع چنین وقایعی است. پس کافی است تعداد این وقایع را به دست آوریم.

به چند طریق می‌توان دو تا H و سه تا T را در ۵ جایگاه قرار داد؟ $\binom{5}{2}$؛

پس داریم:

$$P(B) = \binom{5}{2}p^2q^3$$

که برای $p = \frac{1}{2}$ همان $\frac{\binom{5}{2}}{2^5}$ می‌شود.

حالت کلی این مثال را **آزمایش برنولی** گویند.

آزمایش برنولی (Bernoulli Trials):

واقعهٔ A در فضای Ω را در نظر بگیرید. اگر $P(A)=p$ و $P(\bar{A})=1-p=q$ باشد، در هر بار انجام آزمایش یا A یا $\bar{A}$ اتفاق می‌افتد و آزمایشها در شرایط یکسان تکرار شده و از هم مستقل می‌باشند.

حال اگر واقعهٔ B در فضای $\Omega_n=\Omega\times\cdots\times\Omega$ این باشد که واقعهٔ A، kبار با ترتیب خاصی اتفاق افتد، مثلاً در kبار اول A اتفاق بیفتد (و در $n-k$بار بعدی $\bar{A}$) خواهیم داشت:

$$\Omega_n=\{(00\cdots00),(00\cdots01),\ldots,(11\cdots11)\}$$

$$B=\{(\underbrace{11\ldots1}_{k\text{ بار}}\underbrace{00\ldots0}_{n-k\text{ بار}})\}$$

A_i: مجموعهٔ کلیهٔ nتایی‌های مرتبی که عنصر iام آنها یک است

$$P(B)=P(A_1A_2\cdots A_k\bar{A}_{k+1}\cdots\bar{A}_n)\underset{\text{استقلال آزمایشها}}{=}P(A_1)P(A_2)\cdots P(A_k)P(\bar{A}_{k+1})\cdots P(\bar{A}_n)=\underbrace{pp\cdots p}_{k\text{ بار}}\underbrace{qq\cdots q}_{n-k\text{ بار}}$$

$$=p^kq^{n-k}$$

حال اگر واقعهٔ D در Ω_n این باشد که واقعهٔ A، kبار (با هر ترتیبی) اتفاق افتد، $P(D)=P_n(k)$ چقدر است؟

تعداد وقایع ساده‌ای که در آن A، kبار اتفاق می‌افتد $\binom{n}{k}$ است و همگی احتمال $p^k q^{n-k}$ دارند. پس (طبق اصل ۳) داریم:

$$P_n(k)=\binom{n}{k}p^k q^{n-k}$$

روشن است که:

$$\sum_{k=0}^{n}P_n(k)=1$$

آزمایش برنولی تعمیم‌یافته:

در آزمایش برنولی فقط دو حالت داشتیم: وقوع A یا وقوع $\bar{A}$. در حالت کلی اگر $i = 1, 2, \ldots, r$ و A_iها مجموعهٔ Ω را افراز کنند و احتمال هر یک از آنها p_i $(i = 1, 2, \ldots, r)$ باشد (و آزمایشها در شرایط یکسان تکرار شده و از هم مستقل باشند)، احتمال واقعهٔ B (در Ω_n) که در n آزمایش، A_iها هر یک k_iبار $(k_1 + k_2 + \cdots + k_r = n)$ به ترتیب مشخصی اتفاق افتند برابر است با:

$$p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$$

احتمال واقعهٔ D (در Ω_n) که A_iها هر یک k_iبار (با هر ترتیبی) اتفاق افتند برابر است با:

$$C^{n}_{k_1, k_2, \ldots, k_r} P_1^{k_1} P_2^{k_2} \cdots P_r^{k_r}$$

یعنی:

$$P_n(k_1, k_2, \ldots, k_r) = \frac{n!}{k_1! k_2! \cdots k_r!} p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r} \quad \text{(توزیع چند جمله‌ای)}$$

مثال: جعبه‌ای شامل N ترانزیستور است که Mتا از آنها $(M \leq N)$ خرابند. به طور تصادفی ترانزیستوری را از جعبه برداشته و آزمایش می‌کنیم و دوباره به جعبه برمی‌گردانیم. اگر این کار را nبار انجام دهیم (n انتخاب با جایگزینی) احتمال اینکه kتا از این n ترانزیستور خراب بوده باشند چیست؟

چون ترانزیستور را پس از تست به جعبه برمی‌گردانیم، شرایط آزمایش تغییری نمی‌کند و لذا یک آزمایش برنولی است.

واقعهٔ A مورد نظر در Ω اصلی (تک آزمایش) واقعهٔ خراب بودن ترانزیستور است که داریم:

احتمال خراب بودن: $p = \dfrac{M}{N}$

لذا احتمال این واقعهٔ در Ω_n که kتا از n ترانزیستور خراب باشند (مجموعهٔ کلیهٔ زوجهای مرتبی که k عنصر آنها یک است) برابر است با:

$$p_1 = \binom{n}{k}\left(\frac{M}{N}\right)^k\left(1-\frac{M}{N}\right)^{n-k}$$

اگر انتخاب بدون جایگزینی بود، این مثال را (مثال ۳، فصل ۲، ص ۱۵) قبلاً داشتیم: $\dfrac{\binom{\overset{\text{توپهای قرمز}}{\uparrow}\atop 60}{15}\binom{\overset{\text{توپهای سیاه}}{\uparrow}\atop 40}{5}}{\binom{100}{20}}$.

پس داریم:

$$p_2=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

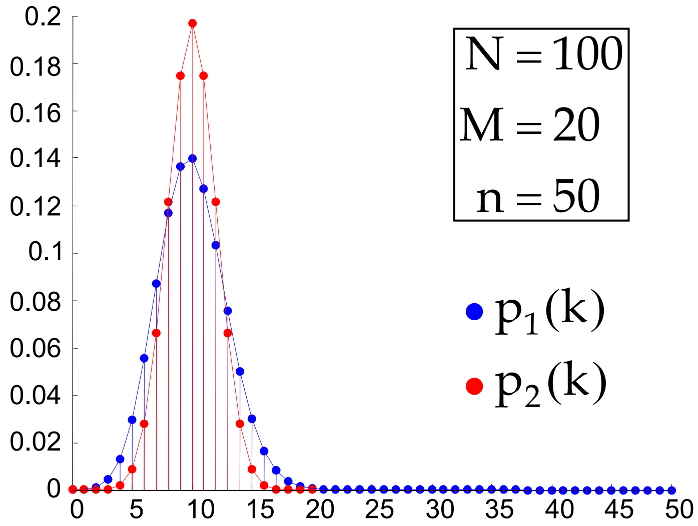
حال اگر بخواهیم مشابه $P_n(k)$ اینها را نیز بر حسب p بیان کنیم، داریم: $p=\frac{M}{N}$ و لذا:

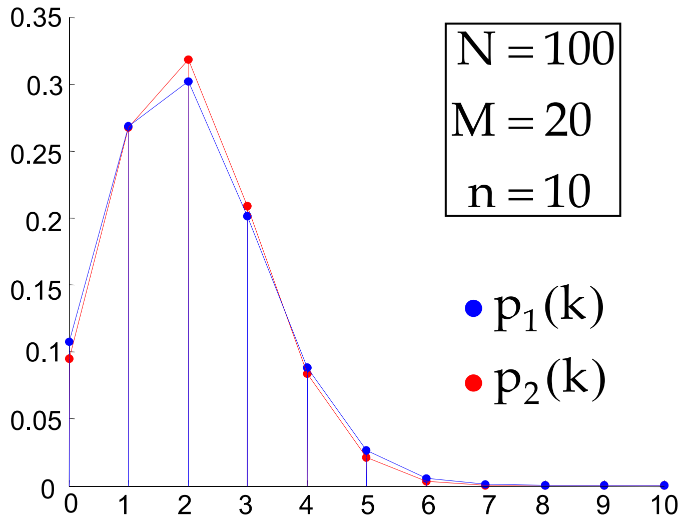
سری فوق هندسی (Hypergeometric): $p_2(k)=\frac{\binom{pN}{k}\binom{N-pN}{n-k}}{\binom{N}{n}}$

در حالی که:

(دوجمله‌ای): $p_1(k)=\binom{n}{k}p^k(1-p)^{n-k}$

اگر $M \ll k$ و $N \ll n$ باشد، داریم: $p_1 \simeq p_2$.





مثال: برای یک آزمایش احتیاج به مقاومت ۱۰ Ω دقیق داریم. تنها $p =$ ٪۱ از مقاومتهای موجود در بازار در بازهٔ مورد نظر ما هستند. چند تا مقاومت بخریم تا به احتمال $P =$ ٪۹۵ مطمئن باشیم که لااقل یکی در بازهٔ مورد نظر ما خواهد بود.

آزمایش تصادفی: انتخاب یک مقاومت ۱۰ Ω از میان مقاومتهای موجود در بازار

Ω: کلیهٔ مقادیر ممکنه برای مقدار یک مقاومت ۱۰ Ω (یا اینکه دو نقطه بگیریم: بودن و نبودن در بازهٔ مورد نظر)

A: مقاومت مورد نظر در بازهٔ مطلوب باشد ($p =$ ٪۱)

Ω_n: کلیهٔ مقادیر ممکنه برای n مقاومت (یا اگر Ω را دو نقطه در نظر بگیریم، Ω_n، 2^n نقطه خواهد داشت)

C: واقعهٔ مورد نظر در Ω_n؛ در n آزمایش لااقل یک بار A اتفاق افتد

حال می‌خواهیم داشته باشیم: $P(C) = P$؛

$$P(C) = \sum_{k=1}^{n} P_n(k) = \sum_{k=1}^{n} \binom{n}{k} p^k q^{n-k} = 1 - P_n(0) = 1 - q^n = 1 - (1-p)^n$$

$$\Rightarrow 1 - (1-p)^n = P \Rightarrow 1 - P = (1-p)^n \Rightarrow n = \frac{\log(1-P)}{\log(1-p)}$$

$$\begin{cases} P = 0.95 \\ p = 0.01 \end{cases} \Rightarrow n = 296$$

مثال: احتمال اینکه نوعی دیود قبل از ۱۰۰۰ ساعت کار خراب شود ۰/۱۵ است. احتمال اینکه در میان ۱۰۰تا از این دیودها، ۹۳تا یا بیشتر قبل از ۱۰۰۰ ساعت سالم باشند (حداکثر ۷تا خراب شوند) چیست؟

Ω: زمان خراب شدن یک دیود

Ω_n: زمانهای خراب شدن n دیود

A: واقعهٔ مورد نظر در Ω؛ خرابی در فاصلهٔ $(0, 1000)$ با احتمال $p =$ ۰/۱۵

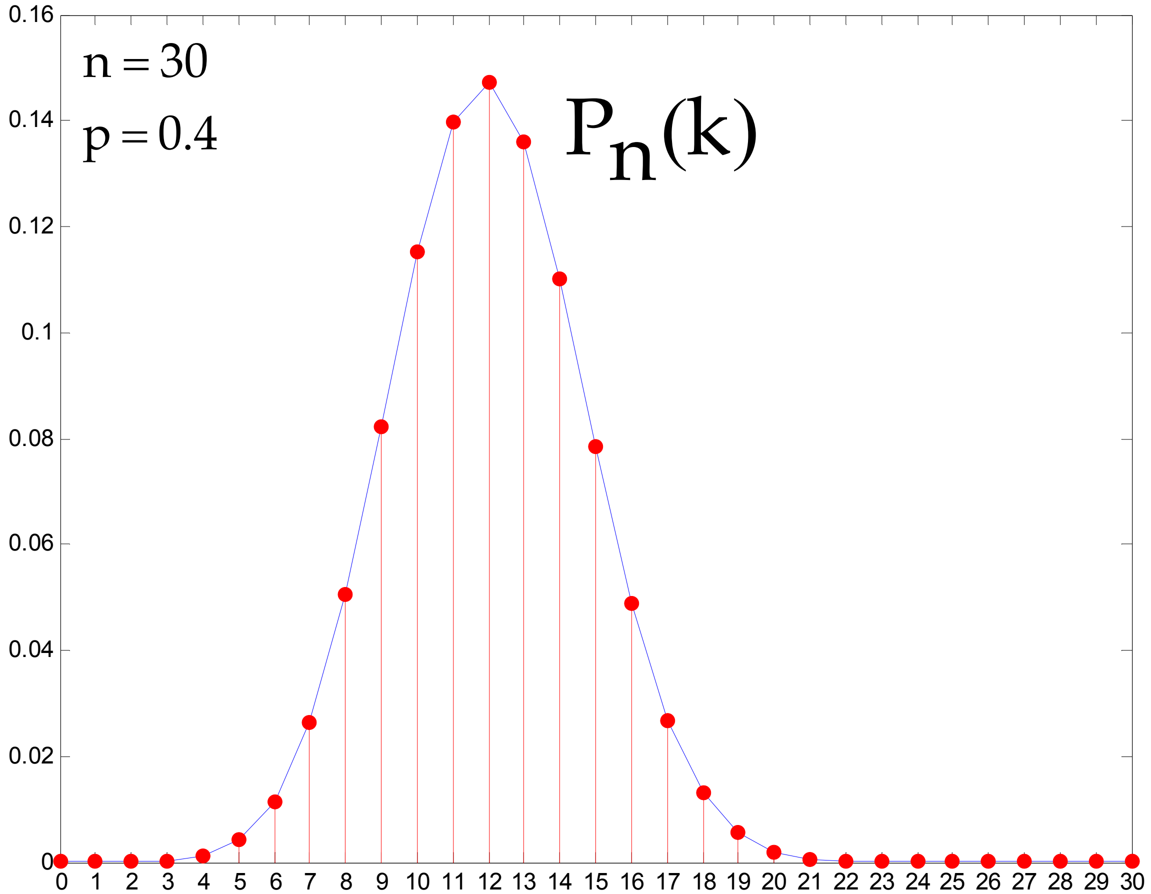
B: واقعهٔ مورد نظر در Ω_n ($n =$ ۱۰۰)؛ حداکثر ۷ دیود خراب شوند (در ۱۰۰ آزمایش، در هفت آزمایش یا کمتر واقعهٔ A اتفاق بیفتد)

$$P(B) = \sum_{k=0}^{7} \binom{100}{k} (0.15)^k (0.85)^{100-k} = 0.012165$$

اگر خرابی یک دیود سبب خرابی (یا افزایش احتمال خرابی) دیود دیگر شود دیگر محاسبهٔ فوق معتبر نخواهد بود. چون آزمایشها وابسته هستند و مقدار احتمال واقعهٔ مورد نظر در یک آزمایش با نتیجهٔ آزمایش قبلی تغییر می‌کند.

مثلاً احتمال اینکه در میان ۹ حرف یک متن انگلیسی دوتا u وجود داشته باشد در حالی که بدانیم کلاً احتمال وقوع u در حروف انگلیسی p است، $\binom{9}{2}p^2q^7$ نخواهد بود. چون بسته به اینکه حرف قبلی چه باشد احتمال u فرق می‌کند. اگر q باشد احتمال وقوع u تقریباً یک خواهد بود. از سوی دیگر اگر حرف قبلی i باشد احتمال وقوع u بسیار کم خواهد بود. یا به همین ترتیب احتمال s به دنبال t یا w زیاد است، ولی احتمال آن پس از j یا f خیلی کم است (رجوع شود به درس رمزنگاری).

برای nهای بزرگ محاسبهٔ $P_n(k)$ مشکل می‌شود و بدتر از آن وقتی است که $\sum$ با تعداد زیاد روی چنین $P_n(k)$هایی داشته باشیم.



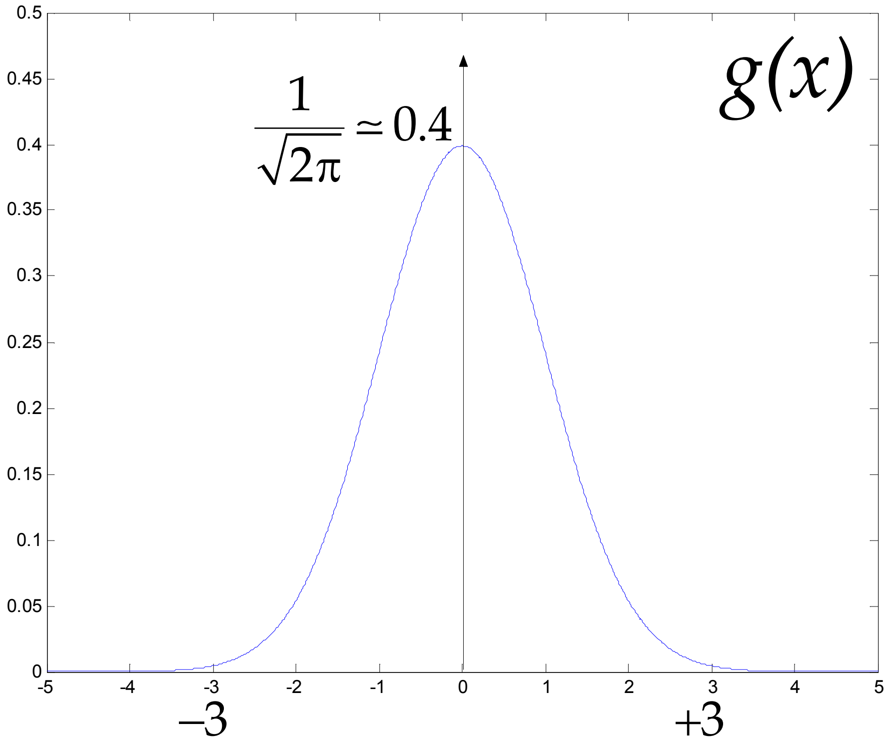
که مقدار ماکزیمم آن در حدود np اتفاق می‌افتد و توضیح دقیق آن در کتاب آمده است.

(اگر $(n+1)p$ صحیح نباشد، $k_m = \lceil (n+1)p \rceil$ و اگر $(n+1)p$ صحیح باشد، $\begin{cases} k_1 = (n+1)p - 1 \\ k_2 = (n+1)p \end{cases}$)

تقریب $P_n(k)$ برای nهای بزرگ:

برای تقریب‌زدن $P_n(k)$ برای nهای بزرگ می‌توان از منحنی نرمال استفاده کرد (شیفت یافته و scale شده).

منحنی نرمال (گوسی):



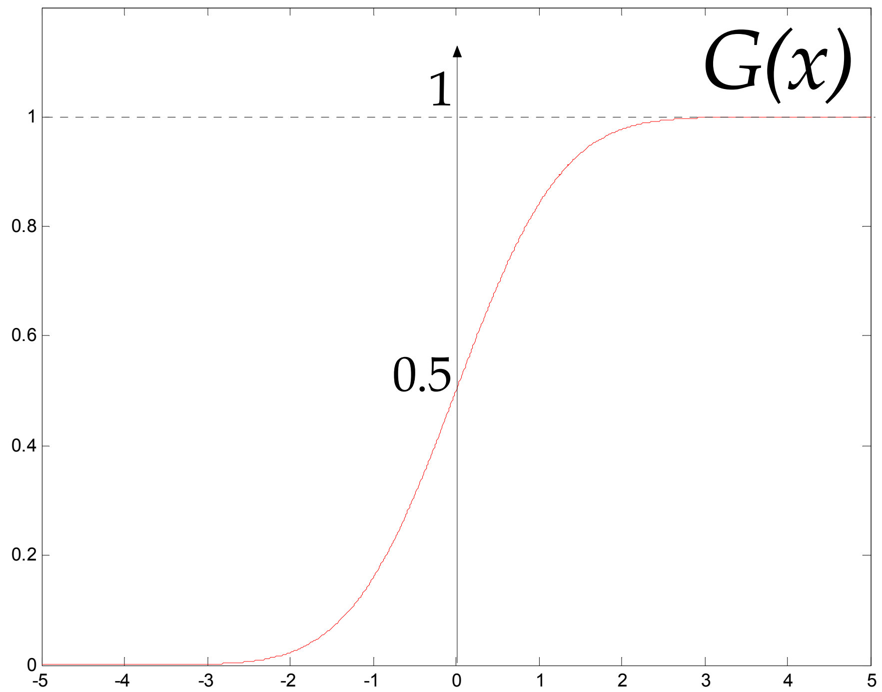
$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$g(x) = g(-x)$$

$$\int_{-\infty}^{+\infty} g(x)dx = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1$$

مقدار $g(x)$ برای خارج بازهٔ $(-3, 3)$ بسیار کوچک است و نقاط عطف آن 1 و 1- هستند.

انتگرال $g(x)$:



$$G(x) = \int_{-\infty}^{x} g(u)du = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

$$G(-\infty) = 0$$

$$G(+\infty) = 1$$

$$G(0) = \tfrac{1}{2}$$

$$G(-x) = 1 - G(x)$$

$G(x)$ در جداول برای $-3 \le x \le 3$ قید شده است.

برای $x > 3$، $G(x)$ خیلی به 1 نزدیک است و داریم:

$$G(x) \simeq 1 - \frac{1}{x} g(x) : x > 3$$

قضیهٔ دموآور-لاپلاس (De Moivre - Laplace Theorem):

برای nهای بزرگ داریم:

$$P_n(k) \simeq \frac{1}{\sigma} g\left(\frac{k-\eta}{\sigma}\right)$$

که در آن: $\begin{cases} \eta = np \\ \sigma = \sqrt{npq} \end{cases}$ است.

یعنی:

$$\binom{n}{k} p^k q^{n-k} \simeq \frac{1}{\sqrt{2\pi}\sqrt{npq}} e^{-\frac{(k-np)^2}{2npq}}$$

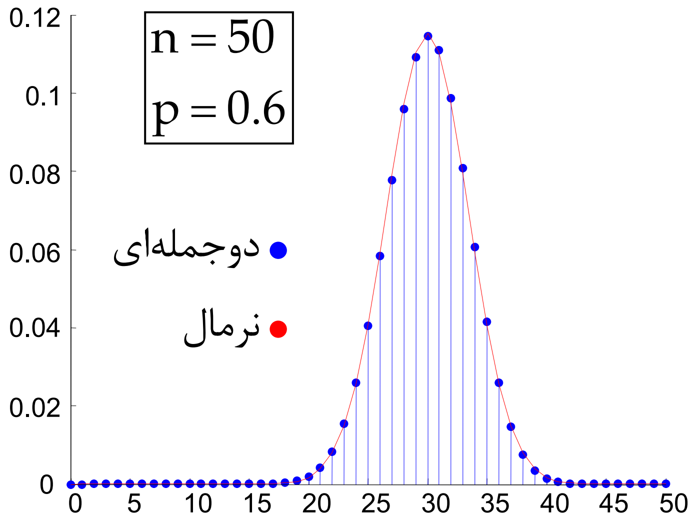
به شرط اینکه: $npq \gg 1$ و $0 < np - 3\sqrt{npq} < k < np + 3\sqrt{npq} < n$.

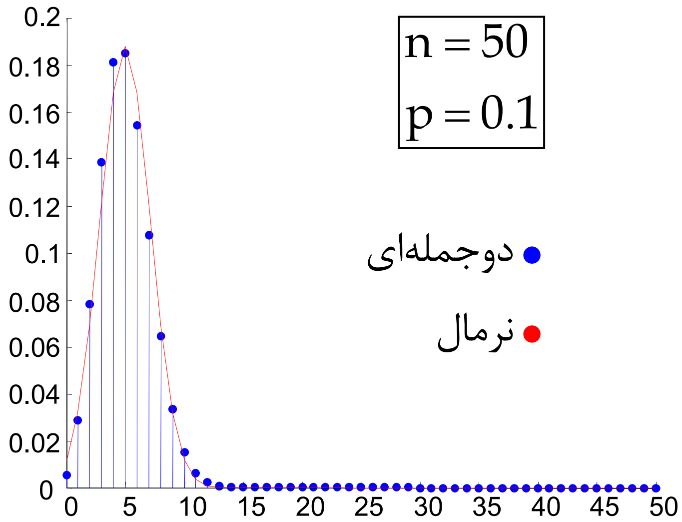
(یعنی کل گسترش ... $P_n(k)$ باید در محدودهٔ 0 و n باشد و نیز k مورد نظر خارج ناحیهٔ 3σ نباشد، چون مقدار تابع ناچیز شده و خطای نسبی تقریب زیاد می‌شود.)

این قضیه را بدون اثبات می‌پذیریم (اثبات در کتاب Beckmann، ص ۳۵).

اصولاً می‌توان نشان داد که حد نسبت اینها یک می‌شود.

این تقریب برای pهای نزدیک به 0.5 بهتر است و هر چه p به صفر یا یک نزدیک شود، تقریب خرابتر می‌شود و n بزرگتری برای خوبی تقریب لازم است. چون تقارن منحنی دوجمله‌ای از بین می‌رود، در حالی که منحنی نرمال متقارن است (بعداً تقریب پواسون را خواهیم دید).





قضیهٔ انتگرال دموآور-لاپلاس:

$$\sum_{k=k_1}^{k_2}\binom{n}{k}p^k q^{n-k} \simeq G\left(\frac{k_2-np}{\sqrt{npq}}\right)-G\left(\frac{k_1-np}{\sqrt{npq}}\right)$$

به شرط اینکه: $(npq \gg 1)$ و $(k_2 < np+3\sqrt{npq} < n)$ یا $(0 < np-3\sqrt{npq} < k_1)$.

(اگر یکی از k_1 یا k_2 خارج از محدوده باشد، جمله‌های متناظر با خارج محدودهٔ 3σ در مقابل سایر جمله‌ها قابل صرف‌نظر است. لذا باز تقریب قابل استفاده است.)

تقریب بهتری هم در کتاب آمده است (ص ۷۴)، یعنی با در نظر گرفتن $k_1 - 0.5$ و $k_2 + 0.5$ که خصوصاً وقتی تفاوت k_1 و k_2 کم باشد، بهتر است.

برای $k_1 = 0$ چون از شرط دوم داریم: $np > 3\sqrt{npq}$، جملهٔ دوم G عددی کوچکتر از 3- می‌شود و لذا تقریباً صفر است.

یعنی:

$$\sum_{k=0}^{k_2}\binom{n}{k}p^k q^{n-k} \simeq G\left(\frac{k_2-np}{\sqrt{npq}}\right)$$

به شرط اینکه: $npq \gg 1$ و $0 < np-3\sqrt{npq} < k_2 < np+3\sqrt{npq} < n$.

مثال: برای مثالی که در صفحهٔ ۱۲ داشتیم (خرابی دیودها)، خواهیم داشت:

$$\begin{cases} p = 0.15 \\ n = 100 \\ k_2 = 7 \end{cases} \rightarrow \sum_{k=0}^{7} P_n(k) = ?$$

$$\begin{cases} np = 15 \\ npq = 12.75 \Rightarrow \sqrt{npq} = 3.57 \Rightarrow \begin{cases} np - 3\sqrt{npq} = 15 - 3(3.57) = 4.3 \\ np + 3\sqrt{npq} = 15 + 3(3.57) = 25.7 \end{cases} \end{cases}$$

$$\rightarrow 12.75 \gg 1 : \boxtimes$$

$$\rightarrow 0 < 4.3 < 7 < 25.7 < 100 : \boxtimes$$

$$\Rightarrow \sum_{k=0}^{7} \binom{100}{7} (0.15)^k (0.85)^{100-k} \simeq G\left(\frac{7-15}{3.57}\right) = G(-2.24) = 0.0125$$

(جواب دقیق 0.0122 بود.)

تعمیم قضیهٔ دموآور-لاپلاس:

$$P_n(k_1,k_2,\ldots,k_r)=\binom{n}{k_1,k_2,\ldots,k_r}p_1^{k_1}p_2^{k_2}\cdots p_r^{k_r}\simeq\frac{e^{-\frac{1}{2}\sum_{i=1}^{r}q_i x_i^2}}{(2\pi)^{\frac{r-1}{2}}\sqrt{p_1p_2\cdots p_r}\sqrt{n^{r-1}}}:x_i=\frac{k_i-np_i}{\sqrt{np_iq_i}}$$

با استفاده از قضیهٔ دموآور-لاپلاس (که حالت خاصی از قضیهٔ حد مرکزی است) می‌توانیم حالت خاصی از **قانون اعداد بزرگ** را که بعداً خواهیم دید اثبات کنیم.

قانون اعداد بزرگ (The Law of Large Numbers):

به وسیلهٔ قضیهٔ لاپلاس دیدیم که اگر واقعهٔ A در n مرتبه تکرار آزمایشی kبار اتفاق افتد، داریم:

$$P\{k_1\le k\le k_2\}\simeq G\left(\frac{k_2-np}{\sqrt{npq}}\right)-G\left(\frac{k_1-np}{\sqrt{npq}}\right)$$

به این وسیله نشان می‌دهیم که برای nهای بزرگ، $\frac{k}{n}$ نزدیک به p خواهد بود. این همان چیزی است که در تعبیر تجربی احتمال داشتیم و حالا مفهوم دقیق ریاضی برای آن بیان می‌کنیم.

قضیه: اگر n_0 به قدر کافی بزرگ انتخاب شود، برای هر $n > n_0$ داریم:

$$P\{|\frac{k}{n} - p| \le \varepsilon\} \ge 1 - \delta$$

به عبارت دیگر برای هر $\varepsilon > 0$ داده شده داریم:

$$P\{|\frac{k}{n} - p| \le \varepsilon\} \xrightarrow{n \to \infty} 1$$

اثبات:

$$P\{|\frac{k}{n} - p| \le \varepsilon\} = P\{(p-\varepsilon)n \le \frac{k}{n} \le (p+\varepsilon)n\} = G(\frac{(p+\varepsilon)n - np}{\sqrt{npq}}) - G(\frac{(p-\varepsilon)n - np}{\sqrt{npq}})$$

$$= G(\varepsilon\sqrt{\frac{n}{pq}}) - G(-\varepsilon\sqrt{\frac{n}{pq}}) = 2G(\varepsilon\sqrt{\frac{n}{pq}}) - 1$$

پس برای هر $\varepsilon > 0$ داده شده داریم: $\varepsilon\sqrt{\frac{n}{pq}} \xrightarrow{n \to \infty} \infty$ و لذا خواهیم داشت: $G(\varepsilon\sqrt{\frac{n}{pq}}) \xrightarrow{n \to \infty} 1$؛

در نتیجه داریم:

$$P\{|\frac{k}{n} - p| \le \varepsilon\} \xrightarrow{n \to \infty} 1$$

مثال: یک سکهٔ سالم را چند بار بیندازیم تا به احتمال ۹۹/۹٪، $\frac{k}{n}$ در بازهٔ (0.49, 0.51) باشد.

$$P\{|\frac{k}{n}-p|\le \varepsilon\}=2G(\varepsilon\sqrt{\frac{n}{pq}})-1$$

$$P\{|\frac{k}{n}-\frac{1}{2}|\le 0.01\}=2G(0.02\sqrt{n})-1=0.999\Rightarrow G(0.02\sqrt{n})=0.9995$$

$$\Rightarrow 0.02\sqrt{n}=3.291\Rightarrow n=27077$$

وقایع نادر:

در آزمایش برنولی اگر واقعهٔ A نادر الوقوع باشد، یعنی $p \ll 1$ باشد، در صورتی که n آن‌قدر بزرگ باشد که $np \gg 1$، می‌توان از قضیهٔ لاپلاس استفاده کرد. ولی اگر np در حدود 1 باشد، شرط صحت قضیهٔ لاپلاس دیگر برقرار نخواهد بود.

قضیهٔ پواسُن (Poisson Theorem):

$$P_n(k) \simeq e^{-\eta} \frac{\eta^k}{k!}$$

یعنی:

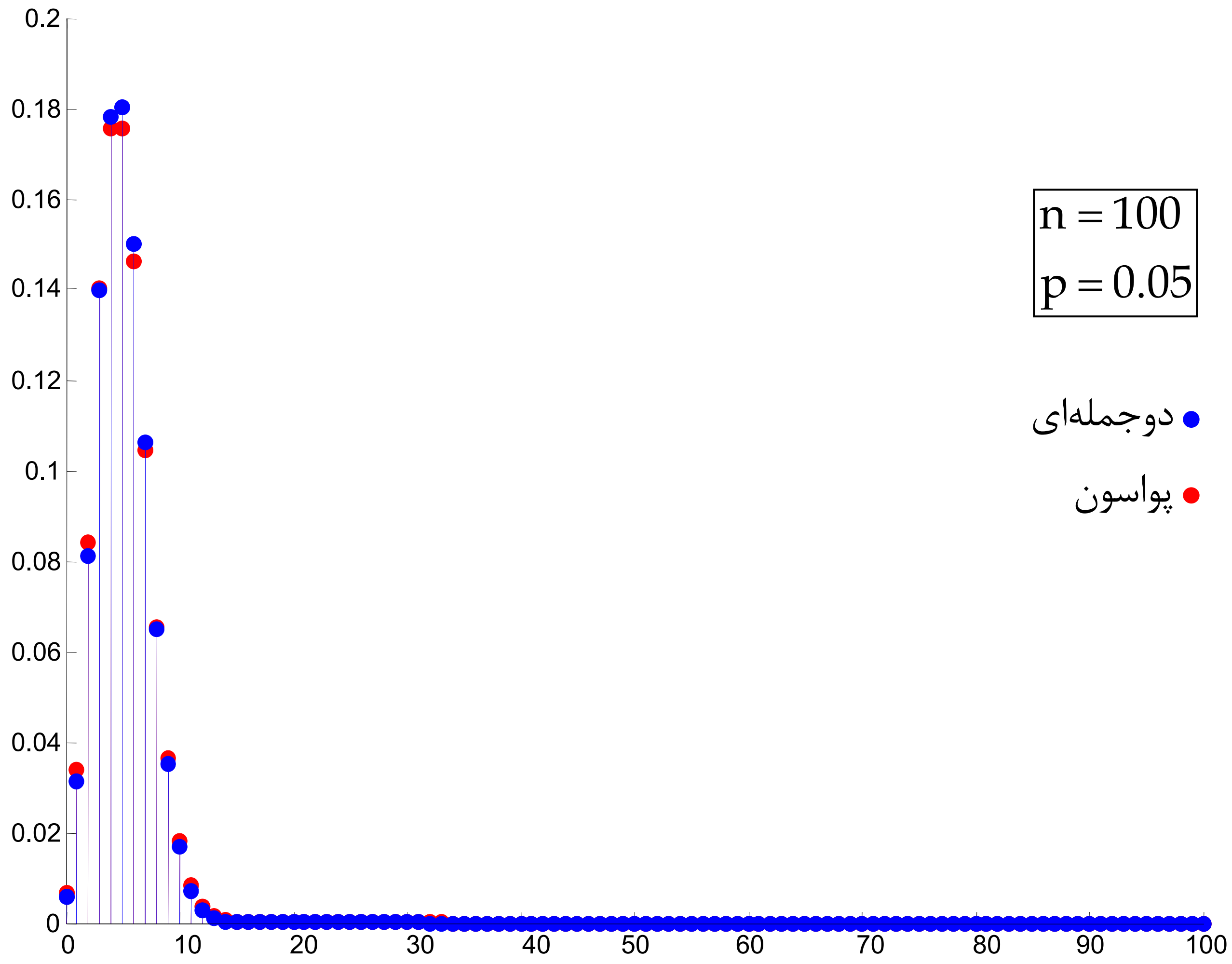
$$\binom{n}{k} p^k q^{n-k} \simeq e^{-np} \frac{(np)^k}{k!}$$

در صورتی که: $p \ll 1$ و $n \gg 1$ و k در حدود np باشد (و np عددی مثلاً در حدود ۱ تا ۱۰ باشد).

(اثبات در کتاب، ص ۷۸)

در واقع می‌توان نشان داد (در کتاب Beckmann، ص ۳۹) که:

$$\lim_{\substack{n \to \infty \\ p \to 0 \\ np \to \eta}} P_n(k) = e^{-\eta} \frac{\eta^k}{k!}$$


n = 100
p = 0.05
دوجمله‌ای
پواسون
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0
10
20
30
40
50
60
70
80
90
100

مثال ۱: در مثالی که قبلاً داشتیم (خرابی دیودها)، $n = 100$ بود. حال اگر $p = 0/03$ باشد، $np = 3$ خواهد بود و داریم:

$$\sum_{k=0}^{7}\binom{100}{k}(0.03)^k(0.97)^{100-k} \simeq \sum_{k=0}^{7} e^{-3}\frac{3^k}{k!} = e^{-3}(1+3+\frac{3^2}{2!}+\frac{3^3}{3!}+\frac{3^4}{4!}+\frac{3^5}{5!}+\frac{3^6}{6!}+\frac{3^7}{7!}) = 0.988$$

جواب دقیق 0.989 می‌باشد.

<u>تعمیم قضیهٔ پواسن:</u>

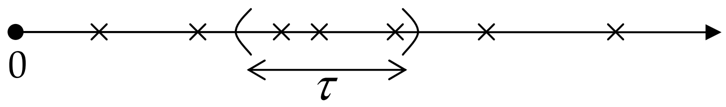
$$P_n(k_1,k_2,\ldots,k_r) = \frac{n!}{k_1!k_2!\cdots k_r!}p_1^{k_1}p_2^{k_2}\cdots p_r^{k_r} \simeq \prod_{i=1}^{r-1} e^{-\eta_i}\frac{\eta_i^{k_i}}{k_i!} : \eta_i = np_i$$

با فرض اینکه برای هر $i = 1,2,\ldots,r-1$ داشته باشیم: $p_i \ll 1$.

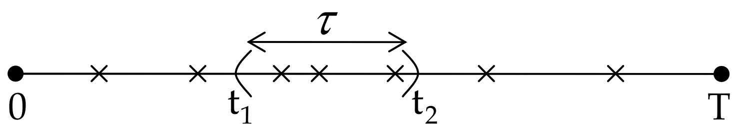
برای η و k بزرگ استفاده از تقریب پواسن مشکل پیدا می‌کند (زیرا $e^{-\eta}\eta^k$ حاصلضرب عددی خیلی کوچک در عددی خیلی بزرگ می‌شود) و می‌توان نشان داد که پواسن را می‌توان با نرمال تقریب زد. یعنی دوباره به استفاده از قضیهٔ لاپلاس برمی‌گردیم (با فرض برقراری شرایط قضیهٔ لاپلاس).

نقاط پواسن:

وقتی با نقاط تصادفی در زمان یا مکان سر و کار داریم توزیع پواسن استفادهٔ زیادی دارد. مثلاً 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 روی یک صفحه، تعداد فوت‌شدگان بیمهٔ عمر در یک سال، تعداد ذرات اتمی منتشر شده از یک مادهٔ رادیواکتیو که به هدف خاصی در طی زمان مشخصی برخورد کنند و یا تعداد الکترونهای منتشر شده از یک مادهٔ فوتوالکتریک در اثر تابش نور در طی مدت زمانی مشخص.



در این آزمایشها: $n \to \infty$ و $p \to 0$، ولی: $np \to \lambda$.



برای n محدود آزمایش تصادفی تکراری ما انتخاب n نقطه در $(0, T)$ است.

یعنی همهٔ n نقطه در فاصلهٔ $(0, T)$ خواهند بود، اما کجای آن را نمی‌دانیم.

احتمال این را می‌خواهیم که kتا از این n نقطه در فاصلهٔ (t_1, t_2) که $t_2 - t_1 = \tau$ است، قرار گیرند. یعنی واقعهٔ $A = \{t \in (t_1, t_2)\}$ در nبار تکرار آزمایش kبار اتفاق افتد.

$$P\{\tau \text{ نقطه در فاصلهٔ زمانی } k\} = \binom{n}{k} p^k q^{n-k} : p = \frac{\tau}{T}$$

اگر n خیلی بزرگ و $p = \frac{\tau}{T}$ خیلی کوچک باشد، با استفاده از تقریب پواسن داریم:

$$P\{\tau \text{ نقطه در فاصلهٔ زمانی } k\} \simeq e^{-\frac{n\tau}{T}} \frac{(\frac{n\tau}{T})^k}{k!} = e^{-\lambda\tau} \frac{(\lambda\tau)^k}{k!} : \lambda = \frac{n}{T} = \text{تعداد نقاط در واحد زمان}$$

اگر $T \to \infty$ (یعنی معادلاً $p \to 0$) و $n \to \infty$، ولی λ را ثابت نگه داریم، تقریب فوق دقیق شده و داریم:

$$P\{\tau \text{ نقطه در فاصلهٔ زمانی } k\} = e^{-\lambda\tau} \frac{(\lambda\tau)^k}{k!}$$

$$P\{\text{ نقطه در واحد زمان } k\} = e^{-\lambda} \frac{\lambda^k}{k!}$$

از جمله خواص نقاط پواسن این است که احتمال وقوع یک نقطه در هر بازه در صورتی که بازه کوچک باشد، متناسب با طول بازه است، زیرا:

$$P\{\text{یک نقطه در فاصلهٔ زمانی } \tau\} = e^{-\lambda\tau}\lambda\tau \overset{\tau \text{ کوچک}}{\simeq} \lambda\tau$$

به عبارت دیگر:

$$\lim_{\tau\to 0}\frac{P\{\text{یک نقطه در فاصلهٔ زمانی } \tau\}}{\tau} = \lim_{\tau\to 0}\lambda e^{-\lambda\tau} = \lambda$$

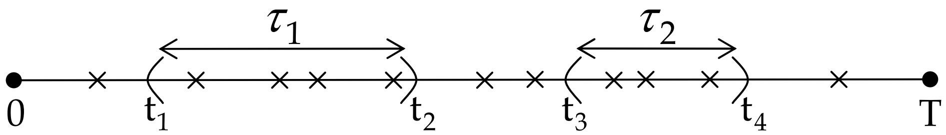
توجه کنید که λ عدد ثابتی است و $\lambda(t)$ نیست. یعنی فرایند ایستان است و نرخ متوسط وقوع پدیده‌ها همیشه یکسان است.

همچنین می‌توان نشان داد که:

$$\lim_{\tau\to 0}\frac{P\{\text{بیش از یک نقطه در فاصلهٔ زمانی } \tau\}}{\tau} = 0$$

(چون احتمال موجود در صورت کسر عبارت است از: $e^{-\lambda\tau}\dfrac{\lambda^2\tau^2}{2!}$)

ویژگی دیگر نقاط پواسن این است که اگر دو بازهٔ غیرمتلاقی را در نظر بگیریم، احتمالها مستقل خواهند بود (اثبات در ص ۸۰ کتاب).



یعنی اگر (t_1, t_2) و (t_3, t_4) دو بازهٔ غیرمتلاقی باشند و $\tau_1 = t_2 - t_1$ و $\tau_2 = t_4 - t_3$ باشد، داریم:

$P\{$ k_1 نقطه در τ_1 و k_2 نقطه در τ_2 $\} = P\{$ k_1 نقطه در τ_1 $\}P\{$ k_2 نقطه در τ_2 $\}$

اصولاً می‌توان (با استفاده از معادلات دیفرانسیل یا با استفاده از توزیع دوجمله‌ای - راس، ص ۱۶۶) نشان داد که اگر بخواهیم تعداد وقوع پدیده‌ای در طی مدت مشخصی از زمان در شرایط زیر صدق کند، توزیع پواسن خواهیم داشت:

۱. احتمال وقوع پدیده در هر فاصلهٔ زمانی در حد (فاصلهٔ زمانی بسیار کوتاه) متناسب با طول فاصله باشد:

$$\mathrm{P}\{\tau \text{ یک نقطه در فاصلهٔ زمانی }\} = \lambda\tau + O(\tau)$$

$O(\tau)$ تابعی است که برای آن $\lim_{\tau\to 0}\frac{O(\tau)}{\tau} = 0$ باشد، یعنی با کاهش τ سریعتر از τ به سمت صفر رود، مانند τ^{α} وقتی که $\alpha > 1$ باشد.

۲. احتمال اینکه پدیده در یک فاصلهٔ زمانی بیش از یک بار اتفاق افتد در حد (فاصلهٔ زمانی کوتاه) صفر باشد:

$$\mathrm{P}\{\tau \text{ بیش از یک نقطه در فاصلهٔ زمانی }\} = O(\tau)$$

۳. تعداد وقوع پدیده در دو بازهٔ زمانی غیرمتلاقی مستقل باشد.

با صرف این سه شرط می‌توان نشان داد که به توزیع پواسن می‌رسیم.

توجه کنید که شرط صحت به کار بردن دوجمله‌ای و تقریبهای آن از جمله پواسن، استقلال آزمایشها بود. البته اگر وابستگی ضعیف باشد نیز تقریب مناسب خواهد بود.

مثلاً در مسألهٔ تطابق، اگر واقعهٔ A_i، قرار گرفتن نامهٔ iام در پاکت خود باشد، داریم:

$$P(A_i)=\frac{1}{n},P(A_iA_j)=\frac{1}{n(n-1)}$$

یعنی A_iها مستقل نیستند. ولی بستگی آنها برای nهای بزرگ خیلی کم است.

لذا منطقی است که تعداد موفقیتها به طور تقریبی دارای توزیع پواسن با پارامتر $np=\frac{n}{n}=1$ باشد.

یعنی برای nهای بزرگ، احتمال هیچ موفقیت برابر با $\frac{1}{e}$ و احتمال k موفقیت برابر با $\frac{e^{-1}}{k!}$ است.