O problema do caminhante aleatório – 1

Alexandre Diehl

Departamento de Física – UFPel

O que é o problema?

The Problem of the Random Walk/Drunkard's Walk

"A man starts from a point 0 and walks l yards in a straight line; he then turns through any angle whatever and walks another l yards in a second straight line. He repeats this process n times. I require the probability that after n of these stretches he is at a distance between r and $r + \delta r$ from his starting point." Karl Pearson. Nature **72**, 294 (1905).

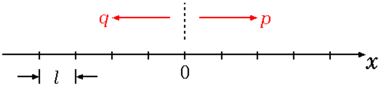
A resposta é dada na semana seguinte por Lord Rayleigh, ao relacionar o problema com vibrações sonoras (1880). Rayleigh propõe que para grandes valores de n, a resposta é dada por

$$\frac{2}{nl^2}\, e^{-r^2/nl^2}\, r\delta r$$

Caraterização do problema

Características dos deslocamentos:

- N passos sucessivos;
- independência estatística;
- mesmo comprimento l;
- probabilidade p para a direita;
- probabilidade q para a esquerda;
- n_1 passos para a direita;
- n_2 passos para a esquerda.



$$p + q = 1$$

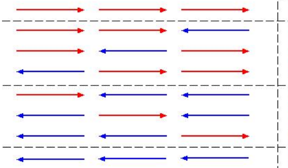
$$N = n_1 + n_2$$

Na versão original, o caminhante executará passos sucessivos, para a direita ou para a esquerda, não podendo ficar parado.

O caminhante aleatório

Caraterização do problema

Caminhante com $N = 3$ passos:



n_1	n_2	m
3	0	3
2	1	1
2	1	1
2	1	1
1	2	-1
1	2	-1
1	2	-1
0	3	-3

Deslocamento líquido,

$$m = n_1 - n_2$$

com

$$-N \leqslant m \leqslant N$$

Posição após N passos: $x = ml$

Neste caso o espaço amostral tem 8 elementos.

De forma geral, temos

$$\frac{N!}{n_1! n_2!}$$

diferentes possibilidades de sequências de N passos, com n_1 deles para a direita e n_2 para a esquerda.

A distribuição de probabilidades

Probabilidade de uma determinada sequência de N passos:

$$P_N(n_1, n_2) = \underbrace{p\,p\ldots p}_{n_1 \text{ fatores}}\ \underbrace{q\,q\ldots q}_{n_2 \text{ fatores}} = p^{n_1}\,q^{n_2}$$

Probabilidade de termos n_1 passos para a direita (e n_2 para a esquerda) após N passos

Distribuição binomial

$$W_N(n_1) = \frac{N!}{n_1!\,n_2!}\,p^{n_1}\,q^{n_2}$$

A distribuição de probabilidades

$$W_N(n_1) = \frac{N!}{n_1!\, n_2!}\, p^{n_1}\, q^{n_2}$$

Como $N = n_1 + n_2 \quad \rightarrow \quad n_2 = N - n_1$

$$W_N(n_1) = \frac{N!}{n_1!\,(N - n_1)!}\, p^{n_1}\, q^{N-n_1}$$

o que corresponde a um dos termos da chamada expansão binomial

$$(p+q)^n = \sum_{x=0}^{n} \left(\begin{array}{c} n \\ \\ x \end{array} \right) p^x\, q^{n-x} = \sum_{x=0}^{n} \frac{n!}{x!\,(n-x)!}\, p^x\, q^{n-x}$$

ou distribuição de Bernoulli, onde p é a probabilidade de ocorrência e q a de não ocorrência.

A distribuição de probabilidades

$$\sum_{n_1=0}^{N} W_N(n_1) = 1 \quad \Longrightarrow \quad \text{Normalização}$$

$$\begin{aligned}\sum_{n_1=0}^{N} W_N(n_1) &= \sum_{n_1=0}^{N} \frac{N!}{n_1!\, n_2!} p^{n_1} q^{n_2} = \sum_{n_1=0}^{N} \frac{N!}{n_1!\,(N-n_1)!} p^{n_1} q^{N-n_1} \\ &= \sum_{n_1=0}^{N} \begin{pmatrix} N \\ n_1 \end{pmatrix} p^{n_1} q^{N-n_1} = \sum_{n_1=0}^{N} (p+q)^N = 1\end{aligned}$$

pois $p + q = 1$ (o caminhante sempre executa um passo).

A distribuição de probabilidades

Deslocamento líquido: $m = n_1 - n_2$ com $N = n_1 + n_2$

$$m = n_1 - n_2 = n_1 - (N - n_1) = 2n_1 - N \quad \rightarrow \quad \begin{cases} n_1 = \frac{N+m}{2} \\ n_2 = \frac{N-m}{2} \end{cases}$$

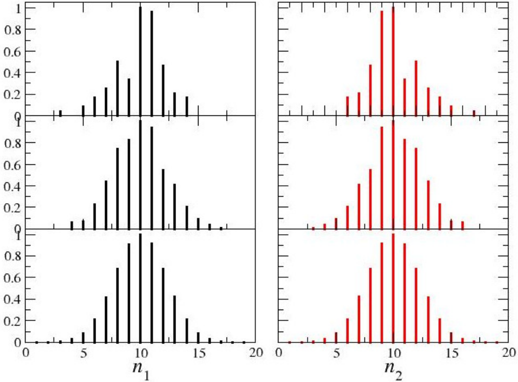
Probabilidade de termos um deslocamento líquido m após N passos

$$P_N(m) = \frac{N!}{\left(\frac{N+m}{2}\right)! \left(\frac{N-m}{2}\right)!} \, p^{\frac{N+m}{2}} \, q^{\frac{N-m}{2}}$$

onde usamos o fato de que $P_N(m) = W_N(n_1)$.

A distribuição de probabilidades

- caminhante aleatório com $N = 20$ passos
- probabilidades para os passos $p = q = 1/2$



Número de repetições:

100

1000

1000000

"Boca de sino" invertida!

Momentos da distribuição: **Valor médio**

$$\langle n_1 \rangle = \frac{\displaystyle\sum_{n_1=0}^{N} n_1\, W(n_1)}{\displaystyle\sum_{n_1=0}^{N} W(n_1)} = Np$$

$$\langle n_2 \rangle = \frac{\displaystyle\sum_{n_2=0}^{N} n_2\, W(n_2)}{\displaystyle\sum_{n_2=0}^{N} W(n_2)} = Nq$$

$$\begin{aligned}
\langle n_1 \rangle &= \sum_{n_1=0}^{N} n_1 \frac{N!}{n_1!\,(N-n_1)!} p^{n_1} q^{N-n_1} = \sum_{n_1=0}^{N} \frac{N!}{n_1!\,(N-n_1)!} \left[p \frac{\partial}{\partial p} \left(p^{n_1}\right) \right] q^{N-n_1} \\
&= p \frac{\partial}{\partial p} \left[\sum_{n_1=0}^{N} \frac{N!}{n_1!\,(N-n_1)!} p^{n_1} q^{N-n_1} \right] = p \frac{\partial}{\partial p} (p+q)^N = pN(p+q)^{N-1} = Np
\end{aligned}$$

onde usamos a condição de normalização para $W_N(n_1)$ e $W_N(n_2)$.

Momentos da distribuição: **Valor médio**

Com isto,

$$\langle n_1 \rangle + \langle n_2 \rangle = N(p+q) = N$$

Em termos do deslocamento líquido $m = n_1 - n_2$

$$\langle m \rangle = \langle n_1 - n_2 \rangle = \langle n_1 \rangle - \langle n_2 \rangle = N(p-q)$$

Para $p = q$,

$$\langle m \rangle = 0$$

Momentos da distribuição: **Variância**

Sinônimos:

- dispersão em relação à média
- segundo momento em torno da média

$$\begin{aligned}
\langle(\Delta n_1)^2\rangle &\equiv \langle(n_1 - \langle n_1\rangle)^2\rangle \\
&= \langle\left(n_1^2 - 2n_1\langle n_1\rangle + \langle n_1\rangle^2\right)\rangle = \langle n_1^2\rangle - 2\langle n_1\rangle^2 + \langle n_1\rangle^2 \\
&= \langle n_1^2\rangle - \langle n_1\rangle^2
\end{aligned}$$

Momentos da distribuição: **Variância**

$$
\begin{aligned}
\langle n_1^2 \rangle &= \sum_{n_1=0}^{N} n_1^2 \, W(n_1) = \sum_{n_1=0}^{N} n_1^2 \, \frac{N!}{n_1! \, (N-n_1)!} \, p^{n_1} \, q^{N-n_1} \\
&= \sum_{n_1=0}^{N} \frac{N!}{n_1! \, (N-n_1)!} \left\{ p \frac{\partial}{\partial p} \left[p \frac{\partial}{\partial p} \, (p^{n_1}) \right] \right\} q^{N-n_1} \\
&= \left(p \frac{\partial}{\partial p} \right) \left\{ \left(p \frac{\partial}{\partial p} \right) \left[\sum_{n_1=0}^{N} \frac{N!}{n_1! \, (N-n_1)!} \, p^{n_1} \, q^{N-n_1} \right] \right\} \\
&= \left(p \frac{\partial}{\partial p} \right) \left[p \frac{\partial}{\partial p} \, (p+q)^N \right] = \left(p \frac{\partial}{\partial p} \right) \left[pN(p+1)^{N-1} \right] \\
&= p \left[N(p+q)^{N-1} + pN(N-1)(p+q)^{N-2} \right]
\end{aligned}
$$

Como $p + q = 1$, $\quad \langle n_1^2 \rangle = Np(1 + Np - p) = Np(Np + q) = (Np)^2 + Npq = \langle n_1 \rangle^2 + Npq$

$$\langle (\Delta n_1)^2 \rangle = Npq$$

Momentos da distribuição: **Desvio Padrão**

A variância nos dá uma ideia da dispersão dos valores de n_1 em torno de seu valor médio.

O desvio padrão nos dá uma medida linear da largura da região sobre a qual os valores de n_1 estão distribuídos.

O desvio padrão é calculado a partir da raiz quadrada da variância:

$$\Delta n_1^\star \equiv \sqrt{\langle(\Delta n_1)^2\rangle} = \sqrt{Npq}$$

Resumo

- Valor médio

$$\langle n_1 \rangle = Np$$

- Variância

$$\langle (\Delta n_1)^2 \rangle = Npq$$

- Desvio padrão

$$\Delta n_1^\star \equiv \sqrt{\langle (\Delta n_1)^2 \rangle} = \sqrt{Npq}$$

Estas quantidades crescem com N

- Desvio relativo

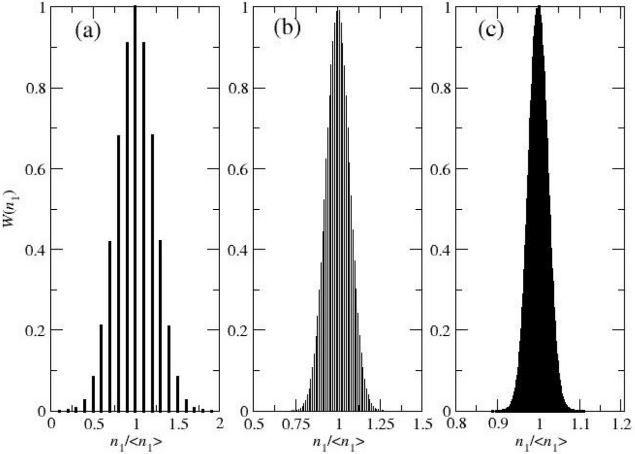
$$\frac{\Delta n_1^\star}{\langle n_1 \rangle} = \left(\frac{q}{p}\right)^{1/2} \frac{1}{\sqrt{N}} ,$$

Esta quantidade decresce como $N^{-1/2}$:

... a distribuição $W_N(n_1)$ torna-se muito fina, centrada em torno de um valor mais provável $\tilde{n}_1$.

Limite assintótico da distribuição

Quando N cresce a distribuição binomial torna-se cada vez mais estreita em torno do valor médio.



- (a) $N = 20$
- (b) $N = 200$
- (c) $N = 2000$

$W_N(n_1)$ torna-se mais estreito a medida que N cresce.

Limite assintótico: distribuição gaussiana

Para $N \to \infty$, $n_1 \to \infty$, tal que em torno do valor mais provável $\tilde{n}_1$,

$$|W(n_1 + 1) - W(n_1)| \ll W(n_1)$$

Quando $N \to \infty$, $W(n_1)$ pode ser tomada como uma função contínua da variável n_1 (que também pode ser considerada contínua nestas condições), próximo ao máximo $n_1 = \tilde{n}_1$.

$$\left.\frac{dW}{dn_1}\right|_{n_1=\tilde{n}_1} = 0 \quad \text{ou} \quad \left.\frac{d \ln W}{dn_1}\right|_{n_1=\tilde{n}_1} = 0$$

Usamos o logaritmo de W porque este varia mais lentamente com n_1, quando comparado com W.

Limite assintótico: distribuição gaussiana

Se tomarmos

$$n_1 = \tilde{n}_1 + \eta \quad \text{(com } \eta \text{ pequeno)}$$

usamos uma série de Taylor em torno de $\tilde{n}_1$,

$$\ln W(n_1) = \ln W(\tilde{n}_1) + B_1\eta + \frac{1}{2!}B_2\eta^2 + \frac{1}{3!}B_3\eta^3 + \ldots$$

onde

$$B_k \equiv \left.\frac{d^k \ln W}{dn_1^k}\right|_{n_1=\tilde{n}_1}$$

$B_1 = 0 \quad$ (extremo) $\qquad B_2 = -|B_2| \quad$ (máximo)

Limite assintótico: distribuição gaussiana

$$\ln W(n_1) = \ln W(\tilde{n}_1) - \frac{1}{2}|B_2|\eta^2 + \frac{1}{6}B_3\eta^3 + \dots$$

$$W(n_1) = \tilde{W}\exp\left(-\frac{1}{2!}|B_2|\eta^2 + \frac{1}{3!}B_3\eta^3 + \dots\right) \qquad \rightarrow \qquad \tilde{W} \equiv W(\tilde{n}_1)$$

Como η é pequeno, mantemos até o termo em segunda ordem

$$W(n_1) = \tilde{W}e^{-\frac{1}{2}|B_2|\eta^2}$$

Distribuição gaussiana

$$W(n_1) = \tilde{W}e^{-\frac{1}{2}|B_2|(n_1-\tilde{n}_1)^2}$$

Limite assintótico: distribuição gaussiana

Significado de $\tilde{n}_1$

Da binomial

$$W_N(n_1) = \frac{N!}{n_1!\,(N-n_1)!}\, p^{n_1}\, q^{N-n_1} \;\rightarrow\; \ln W(n_1) = \ln N! - \ln n_1! - \ln(N-n_1)! + n_1 \ln p + (N-n_1)\ln q$$

Expansão de Stirling ($n \rightarrow \infty$)

$$\ln n! = n \ln n - n + \mathcal{O}(\ln n)$$

$$\begin{aligned} \ln W(n_1) &= N \ln N - N - n_1 \ln n_1 + n_1 - (N-n_1)\,[\ln(N-n_1) - 1] + n_1 \ln p + (N-n_1)\ln q \\ &= N \ln N - n_1 \ln n_1 - (N-n_1)\ln(N-n_1) + n_1 \ln p + (N-n_1)\ln q \end{aligned}$$

$$\begin{aligned} \frac{d \ln W(n_1)}{dn_1} &= -\ln n_1 - 1 + \frac{N}{N-n_1} + \ln(N-n_1) - \frac{n_1}{N-n_1} + \ln p - \ln q \\ &= -\ln n_1 + \ln(N-n_1) + \ln p - \ln q \end{aligned}$$

Limite assintótico: distribuição gaussiana

Significado de $\tilde{n}_1$

$$\left.\frac{d\ln W}{dn_1}\right|_{n_1=\tilde{n}_1} = 0 \quad \Longrightarrow \quad -\ln\tilde{n}_1 + \ln(N-\tilde{n}_1) + \ln p - \ln q = 0$$

$$\ln\left[\frac{N-\tilde{n}_1}{\tilde{n}_1}\,\frac{p}{q}\right] = 0 \quad \Longrightarrow \quad \frac{N-\tilde{n}_1}{\tilde{n}_1}\,\frac{p}{q} = 1 \quad \Longrightarrow \quad (N-\tilde{n}_1)p = \tilde{n}_1 q = \tilde{n}_1(1-p)$$

$$\tilde{n}_1 = Np \quad \Longrightarrow \quad \tilde{n}_1 = \langle n_1 \rangle$$

O valor mais provável é o próprio valor médio.

Limite assintótico: distribuição gaussiana

Significado de B_2

$$\frac{d\ln W(n_1)}{dn_1} = -\ln n_1 + \ln(N - n_1) + \ln p - \ln q$$

$$B_2 = \left.\frac{d^2\ln W}{dn_1^2}\right|_{\tilde{n}_1} = -\frac{1}{\tilde{n}_1} - \frac{1}{N-\tilde{n}_1} = -\frac{1}{Np} - \frac{1}{N-Np} = -\frac{1}{N}\left(\frac{1}{p}+\frac{1}{q}\right) = -\frac{1}{Npq} < 0$$

Como $Npq = \langle(\Delta n_1)^2\rangle \Longrightarrow \quad |B_2| = \dfrac{1}{\langle(\Delta n_1)^2\rangle}$

$|B_2|$ está relacionada com a variância!

$$W(n_1) = \tilde{W}e^{-\frac{1}{2\langle(\Delta n_1)^2\rangle}(n_1-\langle n_1\rangle)^2}$$

Limite assintótico: distribuição gaussiana

Significado de $\tilde{W}$

$$\sum_{n_1=0}^{N} W(n_1) \approx \int W(n_1)dn_1 = \int_{-\infty}^{\infty} W(\tilde{n}_1 + \eta)\, d\eta = 1$$

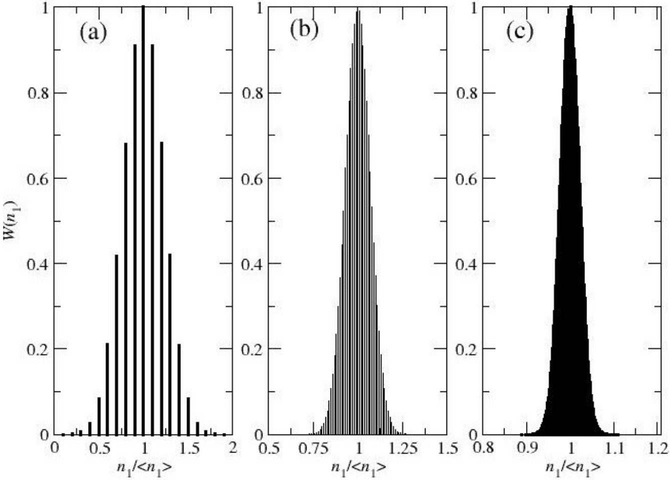
$$\tilde{W} \int_{-\infty}^{\infty} e^{-\frac{1}{2}|B_2|\eta^2}\, d\eta = 1 \quad \Longrightarrow \quad \underbrace{\int_0^{\infty} e^{-\alpha x^2}\, dx = \frac{\sqrt{\pi}}{2}\, \alpha^{-1/2}}_{\text{integral gaussiana}} \quad \Longrightarrow \quad \tilde{W}\, 2\, \frac{\sqrt{\pi}}{2} \sqrt{\frac{2}{|B_2|}} = 1$$

$$\tilde{W} = \sqrt{\frac{|B_2|}{2\pi}} \quad \Longrightarrow \quad \tilde{W} = \sqrt{\frac{1}{2\pi\langle(\Delta n_1)^2\rangle}}$$

Distribuição gaussiana ou normal

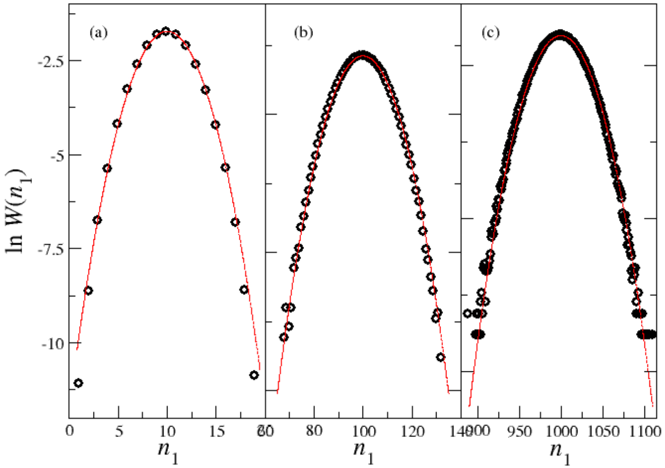
$$W(n_1) = \frac{1}{\sqrt{2\pi\langle(\Delta n_1)^2\rangle}}\, e^{-\frac{1}{2\langle(\Delta n_1)^2\rangle}(n_1 - \langle n_1\rangle)^2}$$

Limite assintótico: distribuição gaussiana



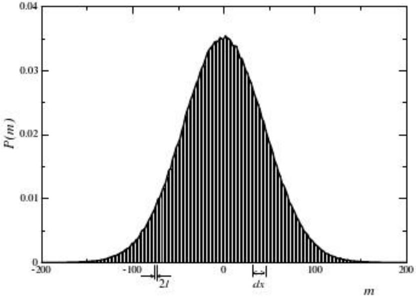
- (a) $N = 20$
- (b) $N = 200$
- (c) $N = 2000$

Limite assintótico: distribuição gaussiana



- (a) $N = 20$
- (b) $N = 200$
- (c) $N = 2000$

Distribuição gaussiana: forma contínua



$$m = 2n_1 - N \quad \Longrightarrow \quad \Delta m = 2$$

$$x = ml \quad \Longrightarrow \quad \Delta x = 2l$$

Variável contínua x:

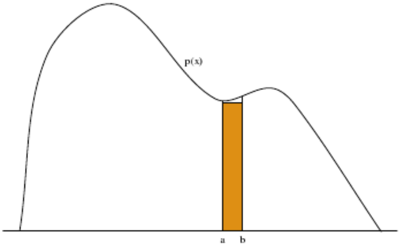
$$\underbrace{\mathcal{P}(x)\,dx}_{\text{probabilidade entre } x \text{ e } x+dx} = P(m)\,\frac{dx}{2l}$$

$\mu \equiv (p-q)Nl \Longrightarrow$ valor médio de x

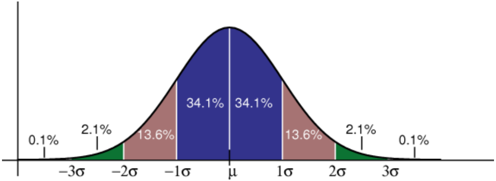
$\sigma \equiv 2\sqrt{Npq}\,l \Longrightarrow$ desvio padrão de x

$$\mathcal{P}(x)\,dx = \frac{1}{\sqrt{2\pi\sigma^2}}\,e^{-(x-\mu)^2/2\sigma^2}\,dx$$

Distribuição gaussiana: forma contínua



$$P(a \;\longrightarrow\; b) = \int_a^b \mathcal{P}(x)\,dx$$



- 68.3% em torno de $\pm\sigma$
- 95.4% em torno de $\pm 2\sigma$
- 99.7% em torno de $\pm 3\sigma$

Teorema do Limite Central

Sejam $x_1, x_2, \ldots$ variáveis aleatórias independentes, identicamente distribuídas (isto é, todas têm a mesma função de probabilidade) e com média μ e variância σ^2 finitas. Então, se

$$s_n = x_1 + x_2 + \ldots + x_n \quad (n = 1, 2, \ldots) \,,$$

teremos como probabilidade

$$\lim_{n\to\infty} P\left(a \leqslant \frac{s_n - n\mu}{\sigma\sqrt{n}} \leqslant b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-u^2/2}\, du \,,$$

isto é, a variável aleatória

$$\frac{(s_n - n\mu)}{\sigma\sqrt{n}} \,,$$

que é a variável padronizada correspondente a s_n, é assintoticamente normal.

Teorema do Limite Central

Em teoria das probabilidades, o teorema do limite central expressa o fato de que qualquer soma de muitas variáveis aleatórias independentes e com mesma distribuição de probabilidade tende a distribuição normal ou Gaussiana.

Como grande parte das características naturais são resultados de diversos fatores, com grande frequência nos deparamos com a distribuição normal.

...

Se um processo aleatório está relacionado com a soma de um número muito grande de processos microscópicos, esta soma estará distribuída de acordo com a distribuição Gaussiana, independentemente da natureza da distribuição dos processos microscópicos.

Limite assintótico: distribuição de Poisson

Eventos Raros

A probabilidade de ocorrência de um evento é pequena, em muitas tentativas de realização.

Suponha que estamos interessados na probabilidade de ocorrência de n eventos, em N tentativas, dado que a probabilidade p do evento em cada tentativa é pequena, ou

$$p \ll 1$$

Distribuição de Poisson

Típica de sistemas onde a probabilidade de ocorrência de um dado evento é muito pequena, tal que o número de vezes que ele ocorre em muitas tentativas é baixa.

Limite assintótico: distribuição de Poisson

Da distribuição binomial,

$$W(n) = \frac{N!}{n!(N-n)!}\, p^n\, (1-p)^{N-n}$$

Como $n \ll N$ e N é grande, usamos a aproximação de Stirling,

$$\begin{aligned} \ln \frac{N!}{(N-n)!} &= \ln N! - \ln(N-n)! \approx N \ln N - (N-n)\ln(N-n) \approx N\ln N - (N-n)\ln N \\ &= N\ln N - N\ln N + n\ln N = n\ln N \end{aligned}$$

$$\frac{N!}{(N-n)!} \approx e^{n\ln N} = N^n$$

Para $p \ll 1$, segue que

$$\ln(1-p)^{N-n} = (N-n)\ln(1-p) = (N-n)\left[-p - \frac{p^2}{2} + \mathcal{O}(p^3)\right] \approx (N-n)[-p] \approx -Np$$

$$(1-p)^{N-n} = e^{-Np}$$

Limite assintótico: distribuição de Poisson

$$W(n) \approx \frac{N^n}{n!}\, p^n\, e^{-Np} = \frac{(Np)^n}{n!}\, e^{-Np}$$

Distribuição de Poisson

$$W(n) \approx \frac{\lambda^n}{n!}\, e^{-\lambda}$$

$$\lambda \equiv Np$$

é o número médio de eventos