Modelo Poisson

Nesse texto apresentamos mais um modelo discreto de distribuição de probabilidades. Os experimentos aleatórios que podem ser estudados por meio desse modelo são aqueles em que se quer estudar o **número de ocorrências de um determinado evento em um intervalo** que pode ser de tempo, de área, de comprimento, de volume, etc.. A grande vantagem desse modelo é que ele fica definido quando se conhece o **número médio de ocorrências do evento nesse intervalo**. Note que, para conhecermos o número médio de ocorrências do evento no intervalo pretendido, basta observamos a ocorrência do evento em qualquer intervalo e determinar a média; como a média é linear, por regra de três obtemos a média no intervalo desejado. Salientamos que, apesar de não ser tratado aqui, um experimento pode ser modelado por Poisson, desde que respeite os Postulados de Poisson (ou Hipóteses de Poisson), apresentados a seguir.

Consideremos uma sequência de intervalos de tamanho **t** unidades, como representado na figura:

Essa sequência de intervalos de tamanho t unidades pode ser considerada uma partição de um eixo ordenado.

H1. **A distribuição do número de ocorrências é a mesma para todos os intervalos da partição**. As variáveis aleatórias, associadas a cada intervalo de tamanho t, têm a mesma distribuição de probabilidades, isto é, o número de ocorrências em um intervalo de tamanho t tem a mesma distribuição de probabilidades que qualquer outro intervalo de mesmo tamanho. Isso significa que o número de ocorrências depende apenas do tamanho do intervalo e não dos extremos.

H2. **Os números de ocorrências registrados nos intervalos de tempo da partição são independentes entre si.** O número de ocorrências, em um intervalo de tamanho t, independe do número de ocorrências em qualquer outro intervalo de mesmo tamanho e não sobreposto. Isso significa que as variáveis aleatórias associadas ao número de ocorrências em intervalos de mesmo tamanho e não sobrepostos são independentes.

H3. Em um intervalo de tamanho pequeno, a probabilidade de se obter uma ocorrência é diretamente proporcional ao tamanho do intervalo. Em um intervalo, suficientemente pequeno, a probabilidade de haver só uma ocorrência é diretamente proporcional ao comprimento do intervalo. Isto é, $P(W_{\Delta t} = 1) = \alpha \Delta t$.

H4. **Em um intervalo de tamanho pequeno, a probabilidade de duas ou mais ocorrências é desprezível.** Em um intervalo suficientemente pequeno, a probabilidade de haver duas ou mais ocorrências é desprezível, isto é, a probabilidade da variável assumir valores maior ou igual a dois pode ser considerada igual a zero $P(W_{\Delta t} \geq 2) = 0$.

H5. **A probabilidade de nenhuma ocorrência em um intervalo de tempo nulo é um.** Chamada condição inicial do modelo, se $\lambda t = 0$ (tamanho do intervalo é zero: $t = 0$), com certeza não teremos ocorrências $P(W_0 = 0) = 1$.

A construção desse Modelo é construída valendo-se desses postulados e não será apresentada neste texto. Nos limitaremos, na apresentação desse modelo, a definir a variável aleatória discreta e a distribuição de probabilidades por meio de uma fórmula.

Seja W a variável aleatória discreta definida por: **"número de ocorrências de um evento em um intervalo de tamanho t"**. Sendo λ o **número médio de ocorrência desse evento no intervalo de tamanho t**, a distribuição de probabilidades, por meio de uma fórmula da variável W, é dada por:

$$P(W = k) = \frac{e^{-\lambda}\,\lambda^k}{k\,!} \quad \text{onde} \quad k = 0, 1, 2, 3, 4, \ldots$$

Observe que a variável pode assumir valores inteiros, 0 , 1 , 2 , 3 , ... até infinito, porque o modelo deve servir para todo tipo de experimento aleatório nessas condições.

O primeiro passo é demonstrar que o modelo está bem definido, isto é, que a soma das probabilidades deve ser igual a 1. De fato,

$$\sum_{k=0}^{\infty} \frac{e^{-\lambda}\,\lambda^k}{k\,!} = \frac{e^{-\lambda}\,\lambda^0}{0\,!} + \frac{e^{-\lambda}\,\lambda^1}{1!} + \frac{e^{-\lambda}\,\lambda^2}{2\,!} + \frac{e^{-\lambda}\,\lambda^3}{3\,!} + \ldots =$$

$$= e^{-\lambda}\left[\frac{\lambda^0}{0\,!} + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2\,!} + \frac{\lambda^3}{3\,!} + \ldots\right] = e^{-\lambda}\;e^{\lambda} = e^{-\lambda+\lambda} = e^0 = 1$$

Na demonstração anterior, utilizamos do Cálculo que a soma $\left[\frac{\lambda^0}{0\,!} + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2\,!} + \frac{\lambda^3}{3\,!} + \ldots\right]$ corresponde ao desenvolvimento em Série de Mac Laurin da função $f(\lambda) = e^{\lambda}$, portanto, nas vizinhanças da origem.

Considerando que o modelo está bem definido, podemos enunciar que a v.a.d. W **"número de ocorrências de um evento em um intervalo de tamanho t"** tem distribuição de Poisson com parâmetro λ (**número médio de ocorrência desse evento no intervalo de tamanho t**) e indicamos: $W \sim P(\lambda)$.

O próximo passo é determinar fórmulas para a média e para a variância, a fim de podermos utilizar nas aplicações.

O valor esperado de W, ou esperança de W, ou média de W é dado por:

$$E(W)=\sum_{k=0}^{\infty} k\,\frac{e^{-\lambda}\lambda^{k}}{k\,!}=0\,\frac{e^{-\lambda}\lambda^{0}}{0\,!}+\sum_{k=1}^{\infty} k\,\frac{e^{-\lambda}\lambda^{k}}{k\,!}=\sum_{k=1}^{\infty} k\,\frac{e^{-\lambda}\lambda^{k}}{k(k-1)!}=$$

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

Logo, $E(W)=\lambda$

Analogamente, como a variância é definida por: $Var(W)=E(W^2)-[E(W)]^2$, temos:

$$E(W^2)=\sum_{k=0}^{\infty}K^2\,\frac{e^{-\lambda}\lambda^{k}}{k!}=0^2\,\frac{e^{-\lambda}\lambda^{0}}{0!}+\sum_{k=1}^{\infty}K^2\,\frac{e^{-\lambda}\lambda^{k}}{k!}=e^{-\lambda}\sum_{k=1}^{\infty}K^2\,\frac{\lambda^{k}}{k!}=e^{-\lambda}\sum_{k=1}^{\infty}K^2\,\frac{\lambda^{k}}{k(k-1)!}=$$

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

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

Portanto, $Var(W)=E(W^2)-[E(W)]^2=\lambda^2+\lambda-\lambda^2=\lambda$. Logo, $Var(W)=\lambda$. Consequentemente, $DP(W)=\sqrt{\lambda}$.

Observe essa característica interessante do Modelo de Poisson de que a média e a variância são iguais ao parâmetro da distribuição. A seguir apresentamos algumas situações para aplicação desse modelo que acabamos de construir.

Exemplo:

Considere que se queira estudar o número de chamadas de emergência que certo Posto de Bombeiros recebe. Sabe-se, de levantamentos anteriores, que tal Posto recebe, em média, 2,6 chamadas de emergência por semana. Determine a probabilidade de o referido Posto receber:

a) 5 chamadas em uma semana;

b) até 3 chamadas em uma semana;

c) pelo menos 4 chamadas em uma semana;

d) 2 chamadas em um único dia;

e) no máximo 3 chamadas em 6 dias;

f) não mais do que 2 chamadas em 10 dias.

Determine, também, o número médio de chamadas recebidas pelo Posto por mês, bem como o desvio padrão do tal número de chamadas mensais.

Inicialmente, devemos identificar que a variável aleatória discreta em estudo, se refere ao número de chamadas que o Posto recebe em algum intervalo de tempo, portanto, número de ocorrências de um determinado evento em um intervalo. Além disso, observamos que a situação fornece a informação do que acontece em média, em relação a essa variável, isto é, que o número médio de chamadas recebidas pelo Posto é, em média, 2,6 chamadas por semana. Essas duas informações permitem que se identifique que a variável segue o Modelo de Poisson.

Para definir a variável aleatória com distribuição de Poisson, como ela depende do tamanho do intervalo, precisamos considerar cada item da situação.

No **item (a)**, como se quer determinar o *número de chamadas que o Posto recebe em uma semana*, o tamanho do intervalo é uma semana.

Teremos, então, que a variável aleatória deve ser definida por: W: **"número de chamadas que o posto recebe em uma semana".**

Assim, $\lambda = 2{,}6$ e, portanto, $\boldsymbol{W \sim P(2,6)}$ (W tem distribuição de Poisson com parâmetro 2,6).

A distribuição de probabilidades de $W \sim P(2{,}6)$ é, então, dada por:

$$P(W = k) = \frac{e^{-2,6}(2{,}6)^k}{k!}, para\ k = 0,1,2,3, \ldots$$

O item (a) pede a probabilidade de 5 chamadas em uma semana. Assim, na distribuição de probabilidades, devemos determinar $P(W = 5)$. Teremos: $P(W = 5) = \frac{e^{-2,6}(2{,}6)^5}{5!} \cong 0{,}073539359$ que, conforme o nosso contrato, terá resposta: a probabilidade de o posto receber 5 chamadas em uma semana é, aproximadamente, $\mathbf{0,073539}$.

Item (b): *até 3 chamadas em uma semana*. Como o tamanho do intervalo é igual ao do item (a), podemos utilizar a mesma variável, ou seja, $W \sim P(2{,}6)$.

Devemos, então, determinar: $P(W \leq 3)$, já que queremos calcular até 3 chamadas em uma semana.

Teremos: $P(W \leq 3) = P(W = 0) + P(W = 1) + P(W = 2) + P(W = 3) =$

$$= \frac{e^{-2,6}(2{,}6)^0}{0!} + \frac{e^{-2,6}(2{,}6)^1}{1!} + \frac{e^{-2,6}(2{,}6)^2}{2!} + \frac{e^{-2,6}(2{,}6)^3}{3!} =$$

$$= e^{-2,6}\left[\frac{(2{,}6)^0}{0!} + \frac{(2{,}6)^1}{1!} + \frac{(2{,}6)^2}{2!} + \frac{(2{,}6)^3}{3!}\right] \cong e^{-2,6}[9{,}909333333] \cong$$

$$\cong 0{,}074273578[9{,}909333333] \cong 0{,}736001644$$

Logo, a probabilidade de o posto receber até 3 chamadas em uma semana é, aproximadamente, $\mathbf{0,736002}$**.**

Item (c): *pelo menos 4 chamadas em uma semana*. Como o tamanho do intervalo é igual ao do item (a), podemos utilizar a mesma variável, ou seja, $W \sim P(2{,}6)$.

Devemos, então, determinar: $P(W \geq 4) = P(W = 4) + P(W = 5) + P(W = 6) + \cdots$ cálculo esse até infinito, dificultando sua determinação. Nesses casos, utilizamos o complementar do evento:

$$P(W \geq 4) = 1 - P(W < 4) = 1 - [P(W = 0) + P(W = 1) + P(W = 2) + P(W = 3)] \cong$$

$$\cong 1 - 0{,}736001644 = 0{,}263998356$$

Logo, a probabilidade de o posto receber pelo menos 4 chamadas em uma semana é, aproximadamente, $\mathbf{0,263998}$.

Item (d): *2 chamadas em um único dia*. Note que quando a situação se refere a 2 chamadas, devemos entender exatamente 2 chamadas. Nesse caso, a variável será: W "número de chamadas que o posto recebe em 1 dia".

Como a média é linear, para determinar o número médio de chamadas em um dia, basta utilizarmos a regra de três. Teremos:

nº de chamadas		nº de dias
2,6	$\longrightarrow$	7
$\boldsymbol{\lambda}$	$\longrightarrow$	1

Dessa forma, $7\lambda = 2{,}6 \Rightarrow \lambda = \frac{2{,}6}{7} \cong 0{,}371428571$. Assim, teremos: $\boldsymbol{\lambda} \cong \mathbf{0,371428571}$, ou seja, em média, esse Posto recebe 0,371428571 chamadas por dia.

Portanto, a variável é tal que: $W \sim P(0{,}371428571)$.

Devemos, então, determinar: $P(W = 2) = \frac{e^{-0{,}371428571}(0{,}371428571)^2}{2!} \cong 0{,}047578554$.

Logo, a probabilidade de o posto receber 2 chamadas em um dia é, aproximadamente, $\mathbf{0,047579}$.

Item (e): *no máximo 3 chamadas em 6 dias*. Nesse caso, a variável será: W "número de chamadas que o posto recebe em 6 dias".

nº de chamadas		nº de dias
2,6	$\longrightarrow$	7
$\boldsymbol{\lambda}$	$\longrightarrow$	6

Dessa forma, $7\lambda = 6 \times 2{,}6 \Rightarrow \lambda = \frac{6\times 2{,}6}{7} \cong 2{,}228571428$. Assim, teremos: $\boldsymbol{\lambda} \cong 2{,}228571428$, ou seja, em média, esse Posto recebe 2,228571428chamadas a cada 6 dias.

Portanto, a variável é tal que: $W \sim P(2{,}228571428)$.

Devemos, então, determinar: $P(W \leq 3) = P(W = 0) + P(W = 1) + P(W = 2) + P(W = 3) =$

$$= \frac{e^{-2{,}228571428}(2{,}228571428)^0}{0!} + \frac{e^{-2{,}228571428}(2{,}228571428)^1}{1!} + \frac{e^{-2{,}228571428}(2{,}228571428)^2}{2!} + \frac{e^{-2{,}228571428}(2{,}228571428)^3}{3!} =$$

$$= e^{-2{,}228571428}\left[\frac{(2{,}228571428)^0}{0!} + \frac{(2{,}228571428)^1}{1!} + \frac{(2{,}228571428)^2}{2!} + \frac{(2{,}228571428)^3}{3!}\right] \cong$$

$$\cong e^{-2{,}228571428}[7{,}556548105] \cong 0{,}107682152[7{,}556548105] \cong 0{,}813705361$$

Logo, a probabilidade de o posto receber no máximo 3 chamadas em 6 dias é, aproximadamente, $\mathbf{0,813705.}$

Item (f): *não mais do que 2 chamadas em 10 dias.* Nesse caso, a variável será: W_{W10} "número de chamadas que o posto recebe em 10 dias".

nº de chamadas		nº de dias
2,6	$\longrightarrow$	7
$\boldsymbol{\lambda}$	$\longrightarrow$	10

Dessa forma, $7\lambda = 10 \times 2{,}6 \Rightarrow \lambda = \frac{10\times 2{,}6}{7} \cong 3{,}714285714$. Assim, teremos: $\boldsymbol{\lambda} \cong 3{,}714285714$, ou seja, em média, esse Posto recebe $3{,}714285714$chamadas a cada 6 dias.

Portanto, a variável é tal que: $W \sim P(3{,}714285714)$.

Devemos, então, determinar: $P(W \leq 2) = P(W = 0) + P(W = 1) + P(W = 2) =$

$$= \frac{e^{-3{,}714285714}(3{,}714285714)^0}{0!} + \frac{e^{-3{,}714285714}(3{,}714285714)^1}{1!} + \frac{e^{-3{,}714285714}(3{,}714285714)^2}{2!} =$$

$$= e^{-3{,}714285714}\left[\frac{(3{,}714285714)^0}{0!} + \frac{(3{,}714285714)^1}{1!} + \frac{(3{,}714285714)^2}{2!}\right] \cong$$

$$\cong e^{-3{,}714285714}[11{,}6122449] \cong 0{,}024372844[11{,}6122449] \cong 0{,}283023434$$

Logo, a probabilidade de o posto receber não mais do que 2 chamadas em 10 dias é, aproximadamente, $\mathbf{0,283023.}$

Determine, também, o número médio de chamadas recebidas pelo Posto por mês, bem como o desvio padrão do tal número de chamadas.

Nesse caso, a variável será: W "número de chamadas que o posto recebe em 30 dias".

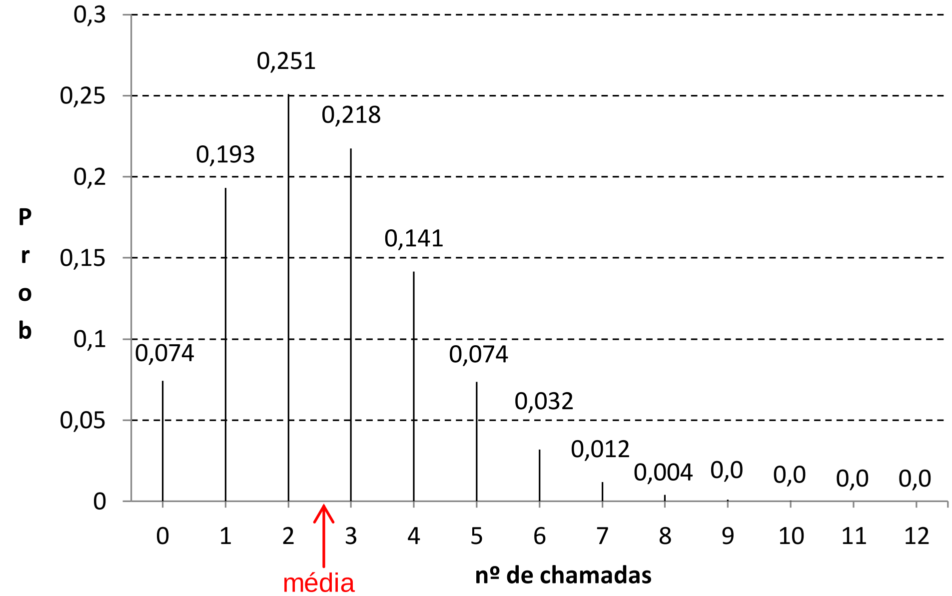
nº de chamadas		nº de dias
2,6	$\longrightarrow$	7
$\boldsymbol{\lambda}$	$\longrightarrow$	**30**0

Dessa forma, $7\lambda = 30 \times 2{,}6 \Rightarrow \lambda = \frac{30\times 2{,}6}{7} \cong 11{,}14285714$. Assim, teremos: $\boldsymbol{\lambda} \cong 11{,}14285714$, ou seja, em média, esse Posto recebe $11{,}14285714$chamadas a cada 6 dias.

Portanto, a variável é tal que: $W \sim P(11{,}14285714)$.

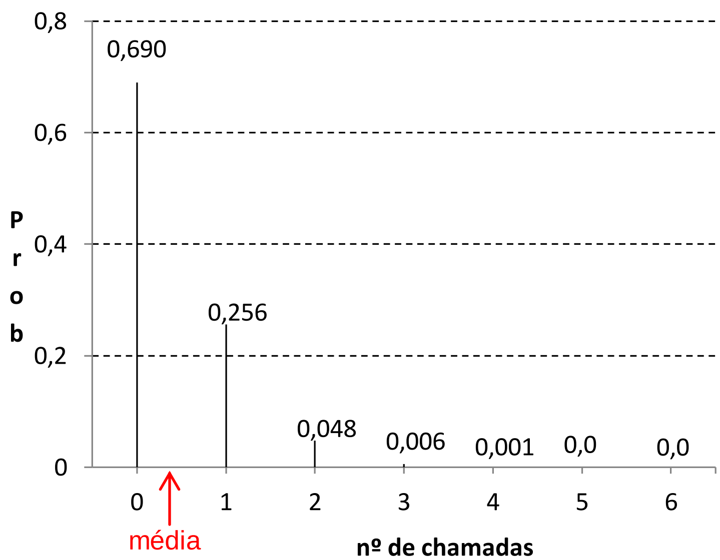
Como $E(W) = \lambda\ e\ Var(W) = \lambda$, o posto recebe, em média, $\mathbf{11,142857}$ chamadas e o desvio padrão é $\sqrt{\mathbf{11,14285714}} \cong \mathbf{3,338092}$ chamadas.

A representação gráfica da variável aleatória $\boldsymbol{W \sim P(2,6)}$ é apresentada a seguir:



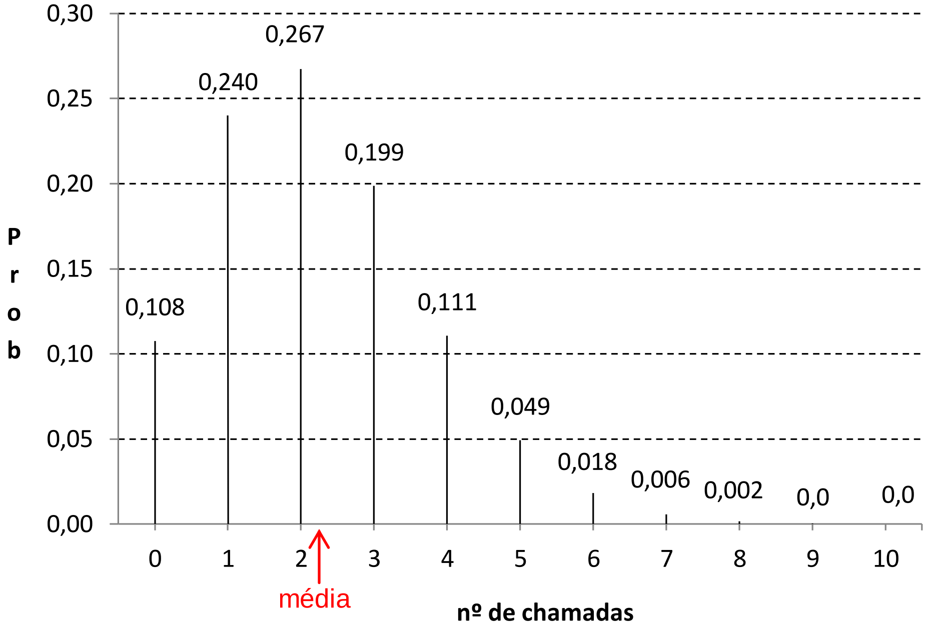
A variável $W \sim P(2{,}6)$ tem média 2,6 chamadas - variância 2,6 (chamadas)^2 e desvio padrão $\sqrt{2{,}6} \cong 1{,}612452$ chamadas.

A representação gráfica da variável aleatória $\boldsymbol{W \sim P(0,371428571)}$ é apresentada a seguir:



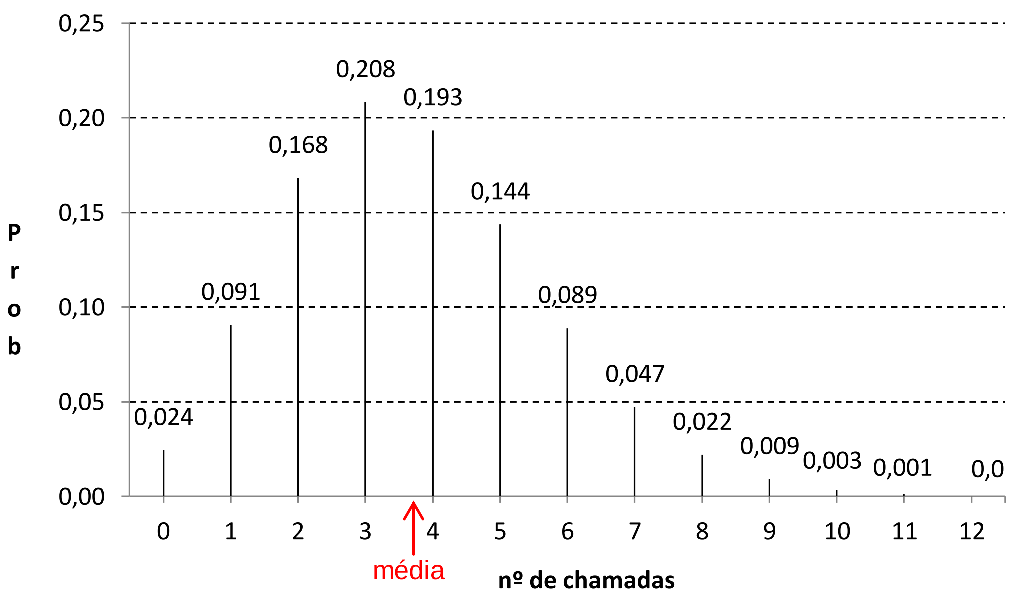
A variável $W \sim P(0{,}371428571)$ tem média 0,371429 chamadas - variância 0,371429 (chamadas)^2 e desvio padrão $\sqrt{0{,}371428571} \cong 0{,}609449$ chamadas.

A representação gráfica da variável aleatória $\boldsymbol{W \sim P(2,228571428)}$ é apresentada a seguir:



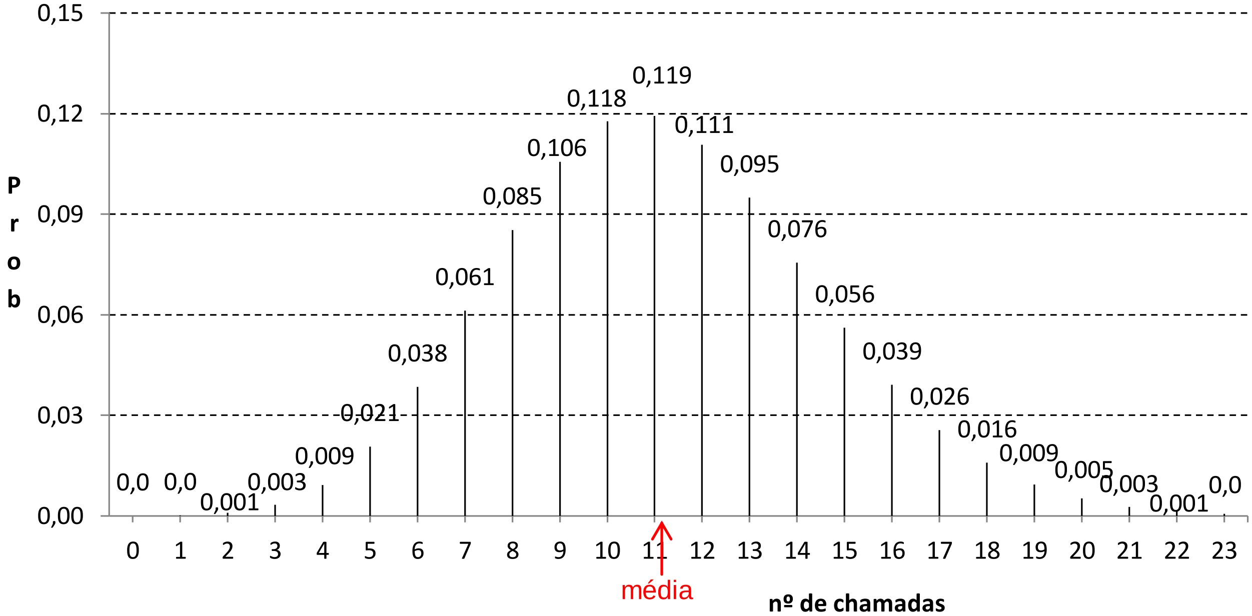
A variável $W \sim P(2{,}228571428)$ tem média 2,228571 chamadas - variância 2,228571 (chamadas)^2 e desvio padrão $\sqrt{2{,}228571428} \cong 1{,}492840$ chamadas.

A representação gráfica da variável aleatória $\boldsymbol{W \sim P(3,7142855714)}$ é apresentada a seguir:



A variável $W \sim P(3{,}7142855714)$ tem média 3,714286 chamadas - variância 3,714286 (chamadas)^2 e desvio padrão $\sqrt{3{,}714285714} \cong 1{,}927248$ chamadas.

A representação gráfica da variável aleatória $\boldsymbol{W \sim P(11,14285714)}$ é apresentada a seguir:



A variável $W \sim P(11{,}14285714)$ tem média $11{,}142857$ chamadas, variância $11{,}142857$ (chamadas)^2 e desvio padrão $\sqrt{11{,}14285714} \cong 3{,}338092$ chamadas.

A seguir apresentamos algumas situações para aplicação do Modelo de Poisson.

EXERCÍCIOS – Modelo de Poisson

1. No estudo do desempenho de uma central de computação, o acesso à Unidade Central de Processamento (CPU) é assumido ser Poisson, com média de 4 requisições por segundo. Essas requisições podem ser de várias naturezas tais como: imprimir um arquivo, efetuar um certo cálculo ou enviar uma mensagem pela Internet, entre outras.

a) Escolhendo-se um intervalo de 1 segundo, qual é a probabilidade de haver mais de 2 acessos à CPU?

b) Considerando um intervalo de 10 segundos, qual a probabilidade de haver 50 acessos?

3. Considere que em um processo de fabricação de placas de vidros, produzem-se pequenas bolhas que se distribuem aleatoriamente pelas placas, com uma densidade média de 0,4 bolha por metro quadrado.

a) Qual o número esperado (média) de bolhas em uma placa de vidro de 3 metros de comprimento por 2 metros de largura?

b) Determine a probabilidade dessa placa não apresentar nenhuma bolha.

c) Outra placa, obtida no mesmo processo, tem dimensões: 1,3m por 2,4m.

(i) Qual o número esperado de bolhas nessa placa?

(ii) Qual a distribuição de probabilidades da v.a. número de bolhas em uma placa de 1,3m por 2,4m?

(iii) Determine a probabilidade de uma dessas placas apresentar pelo menos uma bolha.

(iv) Qual o intervalo da variação do número de bolhas em aproximadamente 99% dessas placas?

(v) Determine a probabilidade de uma dessas placas, escolhida ao acaso, apresentar 300 bolhas.

(vi) Qual o número mais provável de bolhas que se pode ter em uma dessas placas?

4. Em uma estrada há, em média, 2 acidentes a cada 100km. Qual a probabilidade de que em:

a) 250km ocorram pelo menos 3 acidentes?

b) 300km ocorram não mais do que 2 acidentes?

5. Em uma fita de som há, em média, um defeito a cada 200 pés. Qual a probabilidade de que:

a) em 500 pés não aconteça defeito?

b) em 800 pés ocorram pelo menos 3 defeitos?

6. O número de mortes por afogamento em fins de semana, em uma cidade praiana, é, em média, 2 mortes para cada 50.000 habitantes. Qual a probabilidade de que:

a) em 200.000 habitantes ocorram 5 mortes?

b) em 112.500 habitantes ocorram no mínimo 2 mortes?

c) Qual, aproximadamente, o número mais provável de mortes por afogamento, em 112.500 habitantes?

d) Dentre 100.000 habitantes, determine, aproximadamente, o número mínimo e máximo de mortes por afogamento, prováveis de ocorrer.