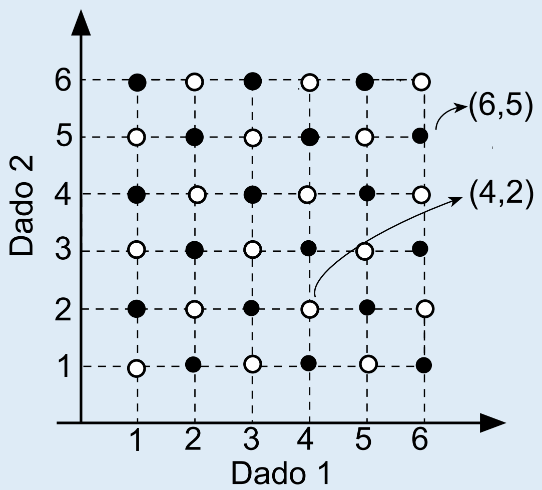
MBA em Ciência de Dados
Aula 3 – Exemplos
Prof. Francisco Rodrigues

1 – Lançamos dois dados e observamos a variável aleatória X é igual a 1 se a som observada for par ou igual a zero, caso contrário. Qual é o valor médio e a variância de X?

RESOLUÇÃO

As saídas do experimento são dadas pelos pontos no gráfico abaixo. Notem que há 36 possíveis saídas.



Os valores possíveis de X são 0 ou 1. Na figura acima, os pontos em preto indicam somas pares e os brancos, ímpares. Logo, a distribuição de X é:

X	0	1
P(X=x)	18/36	18/36

O valor esperado de X:

$$E[X] = \sum_{k=0}^{1} kP(X = k) = 0 \times \frac{18}{36} + 1 \times \frac{18}{36} = \frac{18}{36} = \frac{1}{2}.$$

A variância de X:

$$V(X) = E[X^2] - E[X]^2 = 0^2 \times \frac{18}{36} + 1^2 \times \frac{18}{36} - \left(\frac{1}{2}\right)^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

2 - Uma fábrica produz peças automotivas tais que 5% delas são defeituosas e 95% são não-defeituosas. Se uma peça defeituosa for produzida, o fabricante perde R$ 5,00, enquanto uma peça não-defeituosa lhe dá um lucro de R$10,00. Se *X* for o lucro líquido por peça, calcule *E*(*X*) e V(X).

RESOLUÇÃO

A distribuição de probabilidade do lucro X é dada por:

X	-5	10
P(X=x)	0,05	0,95

Note o que o valor negativo representa um prejuízo de R$5, que é ocorre quando uma peça defeituosa é produzida.
Assim, o valor esperado:

$$E[X] = (-5) \times 0{,}05 + 10 \times 0{,}95 = \ 9{,}25$$

Logo, o lucro médio por peça é igual a R$ 9,25.
No caso da variância, temos:

$$V(X) = E[X^2] - E[X]^2 = (-5)^2 \times 0{,}05 + (10)^2 \times 0{,}95 - (9{,}25)^2 = 10{,}68.$$

Portanto, a variância do lucro é igual a 10,68 reais ao quadrado. Notem que a unidade de medida da variância é a unidade a esperança ao quadrado.

3 - Seja X_i a variável aleatória que é igual a um se sair um sucesso em um experimento aleatório i, ou igual a zero, caso ocorra uma falha. Seja a probabilidade de sucesso $P(X_i = 1) = p$ para $i = 1,2,\ldots,n$. Se a variável aleatória $Y = X_1 + X_2 + \cdots + X_n$ representa o número de sucessos em n experimentos independentes, calcule $E[Y]$ e $V(Y) = \sigma^2$.

RESOLUÇÃO

Para um único experimento i, podemos calcular a esperança de X_i:

$$E[X] = \sum_i x_i P(X = x_i)$$

Como temos dois valores possíveis de X_i, isto é, sucesso e falha, temos:

$$E[X_i] = 0 \times P(X_i = 0) + 1 \times P(X_i = 1) = 0 \times (1-p) + 1 \times p = p$$

Da mesma forma,

$$E\left[X_i^2\right] = \sum_i x_i^2 P(X = x_i) = 0^2 \times (1-p) + 1^2 \times p = p$$

No caso da variável Y:

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

Aplicando a função esperança em ambos os lados, obtemos:

$$E[Y] = E[X_1 + X_2 + \cdots + X_n]$$

Mas a esperança é uma função linear, ou seja, $E[aX + bY + c] = aE[X] + bE[Y] + c$
Logo, a esperança de Y:

$$E[Y] = E[X_1] + E[X_2] + \cdots + E[X_n]$$

Como $E[X_i] = p$, para $i = 1,2,\ldots,n$, temos:

$$E[Y] = p + p + \cdots + p = np.$$

Da mesma forma, a variância de X_i:

$$V(X_i) = E\left[X_i^2\right] - E[X_i]^2 = p - p^2 = p(1-p).$$

Assim, a variância de Y:

$$VY) = V(X_1 + X_2 + \cdots + X_n)$$

Como X_i é independente de X_j para $i \neq j$, temos:

$$V(Y) = V(X_1) + V(X_2) + \cdots + V(X_n) = np(1-p).$$

Assim, temos que $E[Y] = np$ e $V(Y) = np(1-p)$.

4 - O tempo entre falhas em uma linha de produção, representado pela a variável aleatória X, segue o modelo exponencial com $\lambda = 3$ falhas por mês, ou seja,

$$f(x) = \begin{cases} 3e^{-3} \; se \; x \geq 0 \\ 0 \; se \; x < \; 0 \end{cases}$$

Calcule E[X] e V(X).

RESOLUÇÃO

No modelo exponencial, temos:

$$f(x) = \begin{cases} \lambda e^{-\lambda} \; se \; x \geq 0 \\ 0 \; se \; x < \; 0 \end{cases}$$

Sendo a esperança e variância:

$$E[X] = \frac{1}{\lambda}$$

$$V(X) = \frac{1}{\lambda^2}$$

Assim, no caso do tempo entre falhas, temos $\lambda = 3$,

$$E[X] = \frac{1}{3}$$

$$V(X) = \frac{1}{3^2} = \frac{1}{9}$$

5 - Uma fonte radioativa emite partículas alpha a uma taxa igual a $\lambda = 2$ partículas por segundo. Sabendo que o tempo entre emissões segue uma distribuição exponencial, qual é a probabilidade do intervalo entre emissões ser superior ou igual a 7 minutos, sabendo-se que tal intervalo é superior ou igual a 5 minutos?

RESOLUÇÃO

Se X segue o modelo exponencial, então

$$f(x) = \begin{cases} \lambda e^{-\lambda} \; se \; x \geq 0 \\ 0 \; se \; x < \; 0 \end{cases}$$

λ é a taxa com que partículas são emitidas, ou seja, $\lambda = 2$ partículas por segundo.

Queremos calcular:

$$P(X > 7|X > 5) = \frac{P(X > 7, X > 5)}{P(X > 5)}$$

Temos que a interseção dos intervalos X>7 e X>5 é igual a X>7, pois essa é a parte do intervalo eu está nos dois intervalos. Então:

$$P(X > 7|X > 5) = \frac{P(X > 7)}{P(X > 5)}$$

Mas temos que:

$$P(X > t) = \int_t^{\infty} \lambda e^{-\lambda x} dx = \frac{\lambda e^{-\lambda x}}{-\lambda} |_t^{\infty} = e^{-\lambda t}$$

Assim:

$$P(X > 7|X > 5) = \frac{P(X > 7)}{P(X > 5)} = \frac{e^{-7\lambda}}{e^{-5\lambda}} = e^{-2\lambda} = e^{-4} = 0{,}018.$$

Portanto, $P(X > 7|X > 5) = 0{,}018$.

6 - Num livro de 800 páginas há 800 erros de impressão. Qual a probabilidade de que uma página contenha pelo menos 3 erros?

RESOLUÇÃO

Temos: $\lambda = 800/800 = 1$

$$P(X \geq 3) = 1 - P(X < 3) = 1 - [P(X=0) + P(X=1) + P(X=2)] = 1 - 0{,}9196 = 0{,}08.$$

Onde

$$P(X = k)\frac{e^{-\lambda}\lambda^k}{k!}.$$

7 - Numa estrada há 2 acidentes para cada 100 km. Qual a probabilidade de que em:

a) 250 km ocorram pelo menos 3 acidentes?

b) 300 km ocorram 5 acidentes?

RESOLUÇÃO

a) Temos: $\lambda = \frac{250 * 2}{100} = 5\ acidentes\ a\ cada\ 250\ Km.$

$P(X \geq 3) = 1 - P(X < 3) = 1 - [P(X=0) + P(X=1) + P(X=2)] = 1 - 0.124 = 0.87.$

Onde

$$P(X = k)\frac{e^{-\lambda}\lambda^k}{k!}.$$

b) Temos: $\lambda = \frac{300 * 2}{100} = 6\ acidentes\ a\ cada\ 300\ Km.$

$$P(X = 5) = \frac{e^{-6}6^5}{5!} = 0{,}16.$$