Teoria do Risco
Aula3

Danilo Machado Pires
danilo.pires@unifal-mg.edu.br

https://atuaria.github.io/portalhalley/index.html

Função de Distribuição

- Um fenômeno aleatório ou estocástico é descrito minimamente por uma distribuição de probabilidade.
  - Indexa parâmetros e campos de variação

- O conhecimento do modelo e suas principais características permite ao pesquisador ter uma clara visão do uso adequado dos mesmos.

Importantes modelos discretos

Distribuição Uniforme discreta.

$Y \sim U_d(E)$, com "E" sendo o conjunto de seus valores.

$$P(Y = y) = \frac{1}{N} I_{\{1,2,\ldots,N\}}(y)$$

Todos os possíveis valores da variável são equiprováveis.

$Y \sim U_d(1, N)$

$$E(Y) = \frac{N+1}{2} \qquad var(Y) = \frac{N^2-1}{12}$$

Importantes modelos discretos

- **Distribuição de Bernoulli.**

$Y \sim Bernoulli(q)$

$$P(Y = y) = q^y (1 - q)^{1-y} I_{\{0,1\}}(y)$$

Uma variável aleatória que segue o modelo Bernoulli, assume apenas os valores 0 ou 1.

$$E(Y) = q \qquad var(Y) = q(1 - q)$$

Importantes modelos discretos

- **Distribuição Binomial**

Considerando uma sequência de n ensaios de Bernoulli, a observação conjunta de vários desses ensaios leva à definição da distribuição Binomial.

Exemplo:

Suponha o lançamento de 4 moedas, com probabilidade de sair coroa igual a q (sucesso) e $1 - \mathrm{q}$ (fracasso). Qual o modelo de probabilidade para o número de coroas?

Moeda 1	Moeda 2	Moeda 3	Moeda 4	Nº de coroas	Probabilidades	
Cara	Cara	Cara	Cara	0	$q^0(1-q)^4$	$\binom{4}{0} q^0(1-q)^4$
Coroa	Cara	Cara	Cara	1	$q^1(1-q)^3$	$\binom{4}{1} q^1(1-q)^3$
Cara	**Coroa**	Cara	Cara		$q^1(1-q)^3$	
Cara	Cara	**Coroa**	Cara		$q^1(1-q)^3$	
Cara	Cara	Cara	**Coroa**		$q^1(1-q)^3$	
Coroa	**Coroa**	Cara	Cara	2	$q^2(1-q)^2$	$\binom{4}{2} q^2(1-q)^2$
Coroa	Cara	**Coroa**	Cara		$q^2(1-q)^2$	
Coroa	Cara	Cara	**Coroa**		$q^2(1-q)^2$	
Cara	**Coroa**	Cara	**Coroa**		$q^2(1-q)^2$	
Cara	Cara	**Coroa**	**Coroa**		$q^2(1-q)^2$	
Cara	**Coroa**	**Coroa**	Cara		$q^2(1-q)^2$	
Cara	**Coroa**	**Coroa**	**Coroa**	3	$q^3(1-q)^1$	$\binom{4}{3} q^3(1-q)^1$
Coroa	Cara	**Coroa**	**Coroa**		$q^3(1-q)^1$	
Coroa	**Coroa**	Cara	**Coroa**		$q^3(1-q)^1$	
Coroa	**Coroa**	**Coroa**	Cara		$q^3(1-q)^1$	
Coroa	**Coroa**	**Coroa**	**Coroa**	4	$q^4(1-q)^0$	$\binom{4}{4} q^4(1-q)^0$

Importantes modelos discretos

- **Distribuição Binomial**

Seja Y o número total de sucessos obtidos, na realização de n ensaios de Bernoulli independentes. Então é $Y \sim B(n, q)$.

$$P(Y = y) = \binom{n}{y} q^y (1-q)^{n-y} I_{\{0,1,\ldots,n\}}(y)$$

$$E(Y) = nq \qquad var(Y) = nq(1-q)$$

Importantes modelos discretos

- **Distribuição de Poisson**

Sendo a ocorrência do evento em estudo um evento raro, o cálculo através do modelo binomial se torna extremamente laborioso

$Y \sim Poisson(\lambda)$.

$$P(Y = y) = \frac{\lambda^y e^{-\lambda}}{y!} I_{\{0,1,\dots\}}(y)$$

O parâmetro λ indica a taxa de ocorrência por unidade de medida.

$$E(Y) = \lambda \qquad var(Y) = \lambda$$

Exemplo **(Bombas de Londres)**

Durante a Segunda Guerra Mundial, a cidade de **Londres** foi bombardeada intensamente pelos aviões alemães. A discussão sobre a aleatoriedade dos alvos é feita com o auxílio da probabilidade e da estatística. Um possível interesse é saber se houve alguma tendência de concentrar as bombas em alguns alvos, ou se elas foram lançadas aleatoriamente.

..a parte sul da cidade foi repartida em 576 pequenas regiões com meio quilômetro quadrado de área cada uma. Foi contado o número de regiões que receberam k bombas, representado por n_k. O total de bombas na parte sul da cidade foi de 537, resultando numa taxa de $O,9323$ bombas por região.

$$\lambda = \frac{537}{576} = 0{,}9323$$

Exemplo **(Bombas de Londres)**

O modelo de Poisson, com a taxa acima, é considerado adequado para ajustar o número de bombas por região. Isto é, a probabilidade de uma região (hipotética) receber k bombas seria dada por:

$$\boldsymbol{p(k) = \frac{e^{-\lambda}\lambda^{k}}{k!}}$$

Vamos construir uma tabela de frequências esperadas, supondo que o modelo Poisson seja válido, e compará-la com a tabela das frequências observadas.

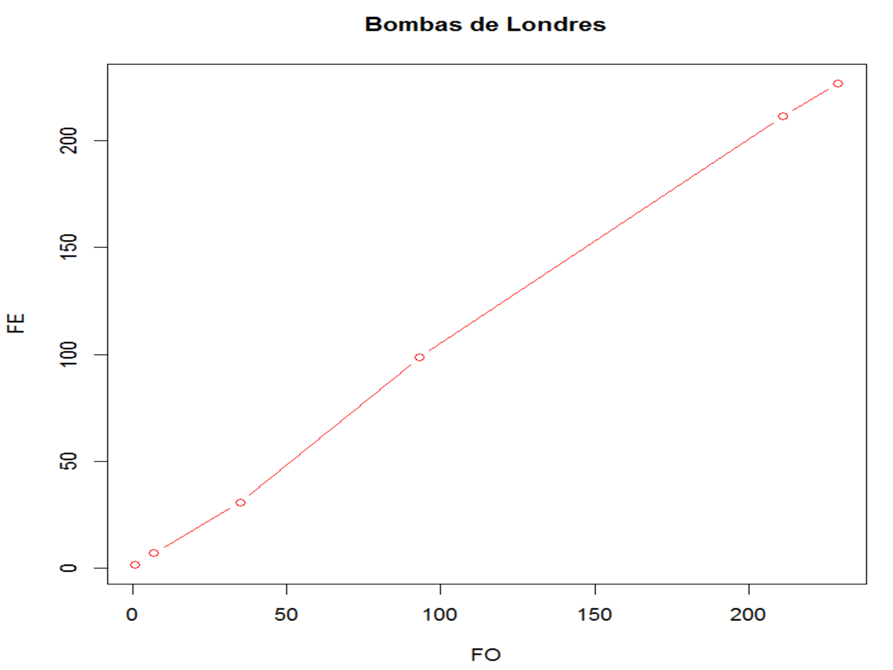
Para obter a frequência esperada do número de regiões com k bombas, denotada por e_k, multiplicamos o total de regiões pela probabilidade $p(k)$. Por exemplo, para a frequência esperada de regiões com 2 bombas, temos

$$\boldsymbol{e_2 = 576\, p(k) = (576)\frac{e^{-0,9323}0,9323^2}{2!} = 98,54}$$

Exemplo **(Bombas de Londres)**

N° Bombas	0	1	2	3	4	5 ou mais
F. observada	229	211	93	35	7	1
F. Esperada	226,74	211,39	98.54	30,62	7,14	1,57

- … os valores observados e esperados estão bastante próximos, servindo de indicação da adequação do modelo proposto.



Exemplo **(Bombas de Londres)**

- .. se houvesse tendência das bombas se agruparem, haveria alta frequência de regiões sem nenhuma bomba, bem como alta frequência de regiões com muitas bombas. Também, teríamos baixa frequência em valores intermediários.

- Pela tabela, percebe-se que isso não ocorre indício que o bombardeio foi feito aleatoriamente sobre essa região.

Importantes modelos discretos

- **Distribuição Geométrica.** $Y \sim G(q)$

$$P(Y = y) = q(1 - q\ )^{y-1}$$

$$E(Y) = \frac{1}{q} \qquad var(Y) = \frac{1-q}{q}$$

- **Distribuição Binomial Negativa** $Y \sim BN(r, p)$

$$P(Y = y) = \binom{y + r - 1}{y} q^r (1 - q\ )^y$$

$$E(Y) = \frac{r(1-q)}{q} \qquad var(Y) = \frac{r(1-q)}{q^2}$$

Importantes modelos contínuos

- **Distribuição Uniforme contínua** $Y \sim U_c(a, b)$

$$\boldsymbol{f(y) = \frac{1}{b-a} I_{[a,b]}(y)}$$

No intervalo $[a, b] \subset \mathbb{R}$, todos os sub-intervalos com mesmo comprimento tem a mesma probabilidade.

$$\boldsymbol{E(Y) = \frac{a+b}{2}} \qquad \boldsymbol{var(Y) = \frac{(b-a)^2}{12}}$$

Importantes modelos contínuos

- **Distribuição Uniforme contínua** $Y \sim U_c[a, b]$.

  O mais simples modelo probabilístico contínuo

$$\boldsymbol{F(y) = \frac{y-a}{b-a} I_{[a,b]}(y) + I_{[b,\infty)}(y).}$$

Importantes modelos contínuos

- **Distribuição Exponencial**

Importante função de distribuição utilizadas na modelagem de dados que representam o tempo até a ocorrência pela primeira vez de algum vento de interesse,

Tempo de falha de um componente eletrônico.

Tempo de ocorrência de indenização em uma seguradora.

Intervalos entre chegadas de mensagens eletrônicas.

Intervalos entre chegadas de chamadas telefônicas a uma central.

Boas propriedades matemáticas.

Importantes modelos contínuos

- **Distribuição Exponencial** $Y \sim Exp(\lambda)$

$$\boldsymbol{f(y) = \lambda e^{-\lambda y} I_{[0,\infty)}(y)}$$

O parâmetro $\boldsymbol{\lambda}$ indica a taxa de ocorrência por unidade de medida, que pode ser tempo, distância ou volume, entre outras.

$$\boldsymbol{F(y) = \left(1 - e^{-\lambda y}\right) I_{(0,\infty)}(y)}$$

$$\boldsymbol{E(Y) = \frac{1}{\lambda}} \qquad \qquad \boldsymbol{var(Y) = \frac{1}{\lambda^2}}$$

Importantes modelos contínuos

- **Distribuição Normal**

→A grande maioria das técnicas empregadas é baseada na distribuição normal.

→Inúmeros fenômenos alheatórios podem ser descritos precisa ou aproximadamente por este modelo.

→Essa distribuição é a forma limitante de outras distribuições de probabilidade, como consequência do teorema centro do limite.

→Muitas estatísticas apresentam normalidade assintótica.

Importantes modelos contínuos

- **Distribuição Normal** $Y \sim N(\mu, \sigma^2)$

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} I_{(-\infty,\infty)}(y)$$

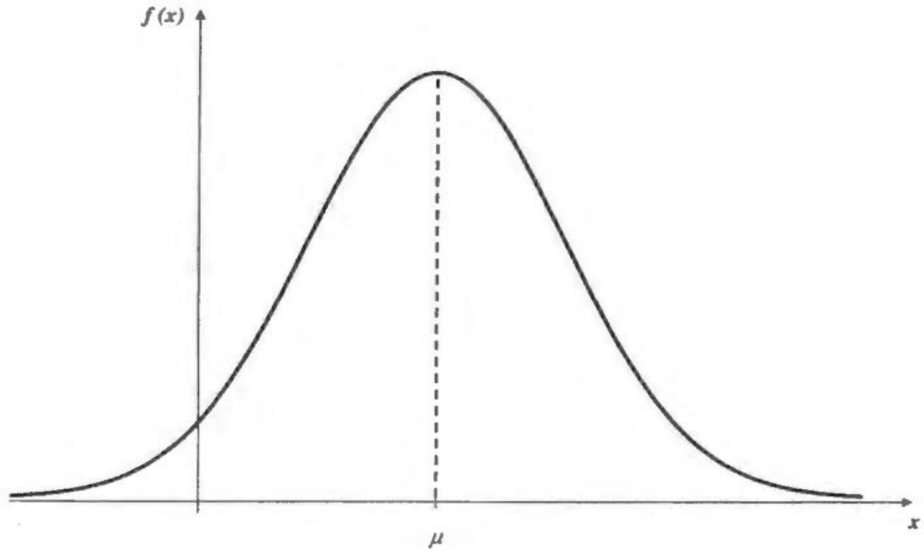
com $\mu, \sigma, y \in \mathbb{R}, \sigma > 0$

Os parâmetros μ, σ^2 são respectivamente, a média e a variância da variável.

$$E(Y) = \mu \qquad var(Y) = \sigma^2$$

Importantes modelos contínuos

- **Distribuição Normal**



Simétrica ao redor de μ e vai diminuindo a massa de probabilidade, à medida que seus valores se movem para as extremidades.

Adequado para várias quantidades envolvendo medidas populacionais:

Peso, Altura, Dosagem De Substâncias No Sangue, Entre Outras.

Importantes modelos contínuos

- A função de distribuição da $\mathrm{N}(\mu, \sigma^2)$ não tem uma forma fechada.
  - Não possui primitiva.

- Os valores de probabilidade são obtidos por integração numérica e apresentados em tabela.

- Basta, tabelar as probabilidades para $\mu = 0$ e $\sigma^2 = 1$. Uma transformação linear de Y é feita nesse sentido.

$$Y = \sigma Z - \mu$$

Sendo $Z \sim N(0,1)$

Importantes modelos contínuos

- **Distribuição Normal**

Sendo $\mathrm{Y} \sim \mathrm{N}(\mu, \sigma^2)$, então $Z = \frac{y-\mu}{\sigma}$ terá distribuição $N(0,1)$.

$$P(Y \leq y) = P\left(Z \leq \frac{y-\mu}{\sigma}\right) = \Phi(z)$$

A distribuição $N(0,1)$ é denominada Normal Padrão ou Normal Reduzida.

Importantes modelos contínuos

- **Distribuição Normal**

$$P(a < Y < b) = P\left(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right)$$

$$P(a < Y < b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

Importantes modelos contínuos

- **Pareto** $X \sim Pareto(\alpha, \beta)$

$$f(X) = \frac{\alpha\beta^{\alpha}}{(\beta+x)^{\alpha+1}}, \quad x > 0 \ (\alpha, \beta > 0)$$

$$E(X) = \frac{\beta}{(\alpha - 1)} \qquad var(X) = \frac{\alpha\beta^2}{(\alpha-1)^2(\alpha-2)}, \alpha > 2$$

- **Lognormal** $Y \sim LN(\mu, \beta)$

$$f_Y(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\log(y)-\mu)^2}{2y\sigma^2}} I_{(0,\infty)}(y)$$

$$E(Y) = e^{\mu+\left(\frac{\sigma^2}{2}\right)} \qquad var(Y) = e^{2(\mu+\sigma^2)}$$