Estatística II

Aula 5

Nelson Faustino [1]

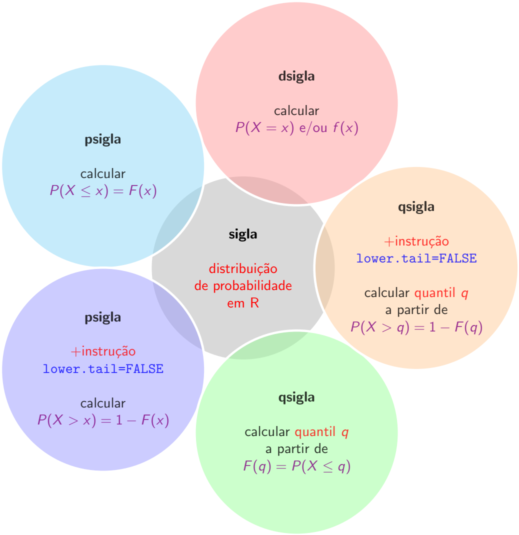
[1] **Faculdade de Economia (FEUC)**
Universidade de Coimbra – Portugal

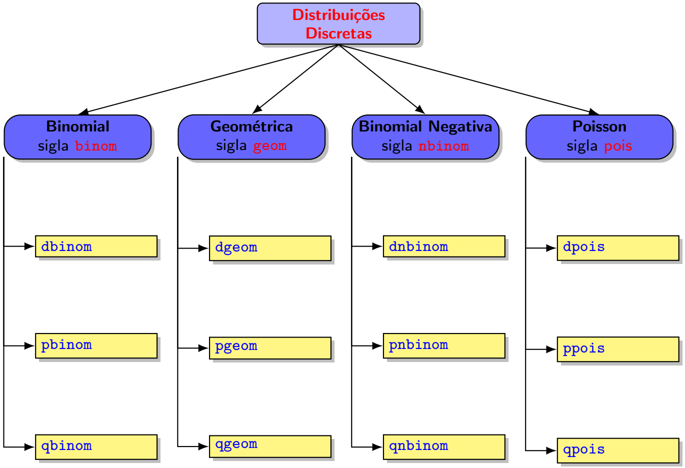
nelson@fe.uc.pt

Conteúdos

Distribuições em R

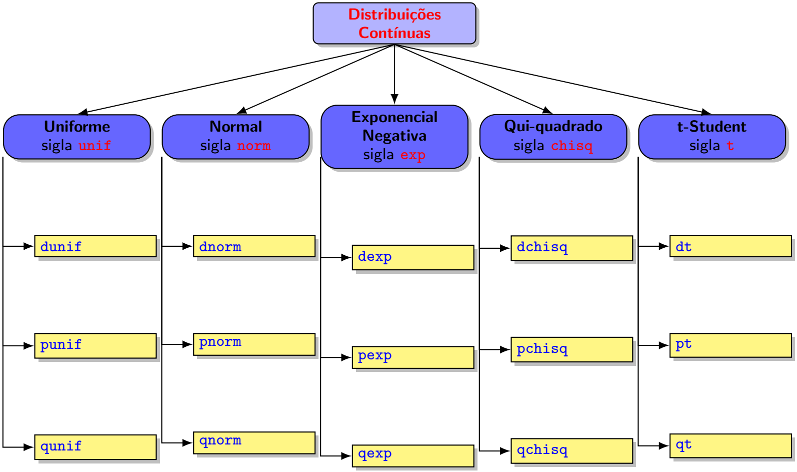
Resumo da Síntaxe




Distribuições Discretas
Binomial
sigla binom
Geométrica
sigla geom
Binomial Negativa
sigla nbinom
Poisson
sigla pois
dbinom
pbinom
qbinom
dgeom
pgeom
qgeom
dnbinom
pnbinom
qnbinom
dpois
ppois
qpois

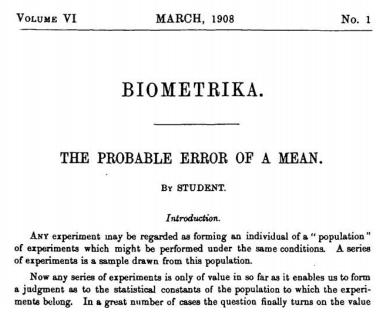
Distribuições em R

Distribuições Contínuas



Curiosidade

t-Student



VOLUME VI MARCH, 1908 No. 1

BIOMETRIKA.

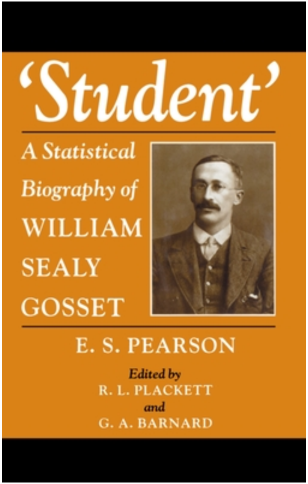
THE PROBABLE ERROR OF A MEAN.

BY STUDENT.

Introduction.

ANY experiment may be regarded as forming an individual of a "population" of experiments which might be performed under the same conditions. A series of experiments is a sample drawn from this population.

Now any series of experiments is only of value in so far as it enables us to form a judgment as to the statistical constants of the population to which the experiments belong. In a great number of cases the question finally turns on the value

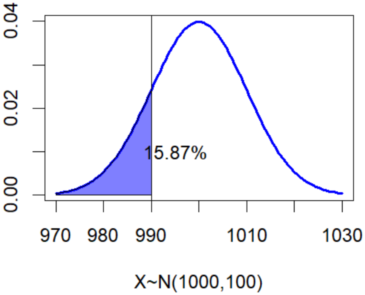


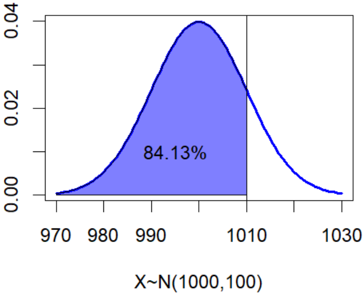
Quem foi William Sealy Gosset?

N. Faustino (Dezembro 9, 2021)
De Mestre Cervejeiro a Estatístico Incógnito, jornal 'O Alcoa'.

Conteúdos

Ilustração gráfica da igualdade $P(X > 1010) = P(X < 990)$:





Distribuição Normal $N(\mu, \sigma^2)$

Cálculo de Probabilidades

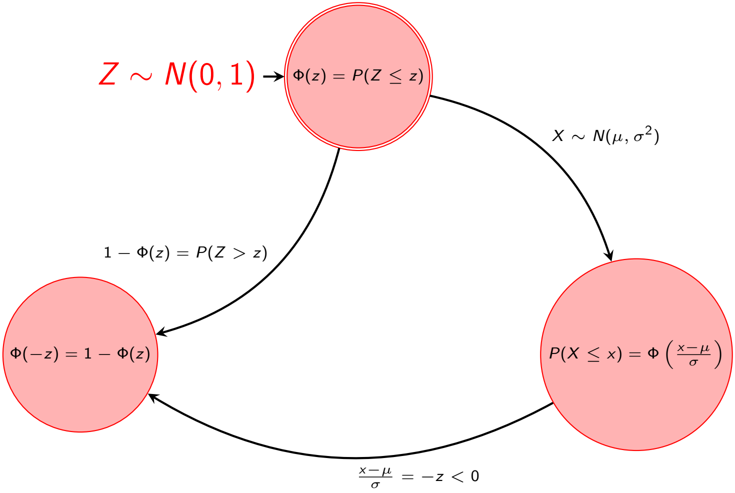
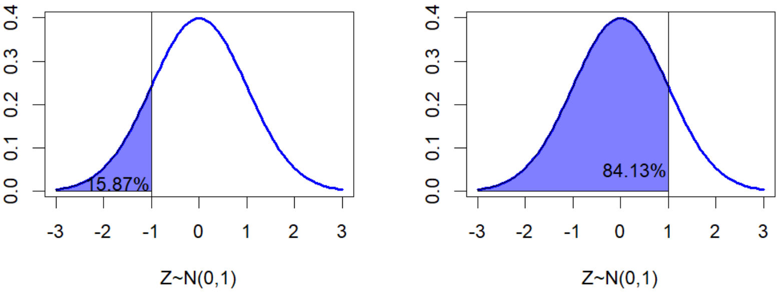
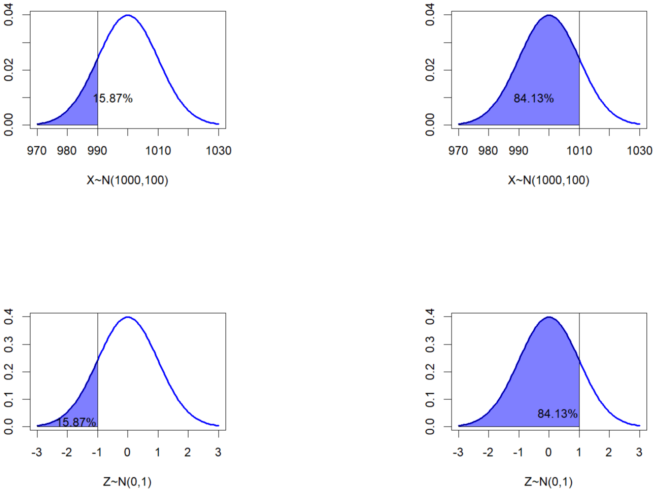


Ilustração gráfica da igualdade $\Phi(-1) = 1 - \Phi(1)$:



Distribuição Normal $N(\mu, \sigma^2)$

Estandartização $Z = \frac{X-1000}{10} \sim N(0,1)$



Distribuição Exponencial Negativa

Probabilidade de **tempo de espera entre acontecimentos consecutivos**

Função densidade de probabilidade ($X \sim Ex(\lambda)$)

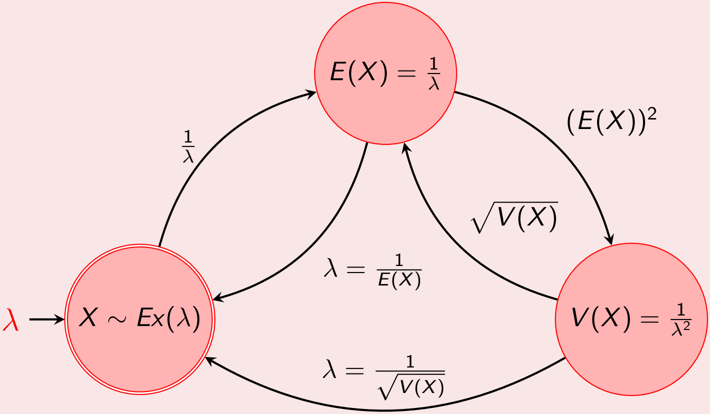
$$f(x) = \begin{cases} \lambda e^{-\lambda x} & (x > 0) \\ 0 & (\text{outros } x) \end{cases}$$

Principais características

- **FGM:** $M(s) = \left(1 - \frac{s}{\lambda}\right)^{-1}$ $(s < \lambda)$.
- **Média:** $E(X) = \frac{1}{\lambda}$, sendo $\dfrac{1}{\lambda} = M'(0)$.
- **Variância:** $V(X) = \frac{1}{\lambda^2}$, sendo $\dfrac{1}{\lambda^2} = M''(0) - (M'(0))^2$.

Distribuição Exponencial Negativa

Cálculo de $E(X)$ e $V(X)$ a partir de λ, e vice-versa



Função de Distribuição $(X \sim Ex(\lambda))$

$$F(x) = \begin{cases} 0 & (x < 0) \\ 1 - e^{-\lambda x} & (x \geq 0) \end{cases}$$

Quantil de ordem α (q_α)

$$F(q_\alpha) = \alpha \quad \Leftrightarrow \quad e^{-\lambda q_\alpha} = 1 - \alpha$$
$$\Leftrightarrow \quad q_\alpha = \frac{-\ln(1-\alpha)}{\lambda}$$

[Propriedade condicional da] "falta de memória" de $X \sim Ex(\lambda)$

$P(X > x + h | X > x) = P(X > h) \;\; (x, h > 0)$, uma vez que

$$\begin{aligned} P(X > x + h | X > x) &= \frac{\overbrace{P(X > x + h \wedge X > x)}^{=P(X>x+h)}}{P(X > x)} \\ &= \frac{1 - F(x+h)}{1 - F(x)} = \frac{e^{-\lambda(\cancel{x}+h)}}{\cancel{e^{-\lambda x}}} = P(X > h). \end{aligned}$$

Função **pexp**

Síntaxe

Comandos R para calcular função distribuição exponencial negativa:

- **pexp(t)** – Calcula, por defeito, $P(T \leq t) = F(t)$ ($T \sim Ex(1)$);
- **pexp(t,lower.tail = FALSE)** – Calcula, por defeito, $P(T > t) = 1 - F(t)$ ($T \sim Ex(1)$);
- **pexp(t,rate)** – Calcula $P(T \leq t) = F(t)$ ($\lambda =$**rate**) para $T \sim Ex(\lambda)$.
- **pexp(t,rate,lower.tail = FALSE)** – Calcula $P(T > t) = 1 - F(t)$ ($\lambda =$**rate**) para $T \sim Ex(\lambda)$.

Exercícios Resolvidos

Exercício 48., página 336

Dados do Exercício 48.a)

Distribuição de Probabilidade: $X \sim N(100, 225)$ de média e variância conhecida ($\mu = 100$ & $\sigma^2 = 225$).

Resolução 48.a)

Em termos da função distribuição de probabilidade $\Phi(z) = P(Z \leq z)$ ($Z = \frac{X-100}{\sqrt{225}} = \frac{X-100}{15} \sim N(0,1)$), a probabilidade pedida é dada por

$$\begin{aligned} P(80 \leq X \leq 115) &= P(X \leq 115) - P(X \leq 80) \\ &= \Phi\left(\frac{115-100}{\sqrt{225}}\right) - \Phi\left(\frac{80-100}{\sqrt{225}}\right) \\ &= \Phi(1) - \Phi\left(-\frac{4}{3}\right) \\ &= \Phi(1) - 1 + \Phi\left(\frac{4}{3}\right) \end{aligned}$$

Exercícios Resolvidos

Exercício 48., página 336

48.a) via > `pnorm(q)`

Duas possíveis resoluções:

```
> pnorm(1)-1+pnorm(4/3)
[1] 0.7501335
> pnorm(1)-pnorm(-4/3)
[1] 0.7501335
```

48.a) via > `pnorm(q,mean,sd)`

```
> pnorm(115,100,15)-pnorm(80,100,15)
[1] 0.7501335
```

Probabilidade a calcular em 48.b)

Corresponde à probabilidade condicional $P(X < 115|X > \mathrm{Me})$, onde Me corresponde ao valor da mediana de $X \sim N(100, 225)$.

Obs: Me $= \mu = 100$, por ser uma distribuição simétrica relativamente à média, i.e. $P(X \leq 100) = 0.5$.

Resolução 48.b)

Usando a definição de probabilidade condicional, segue que

$$\begin{aligned} P(X < 115|X > \mathrm{Me}) &= \frac{P(\mathrm{Me} < X < 115)}{P(X > \mathrm{Me})} \\ &= \frac{P(X < 115) - P(X \leq \mathrm{Me})}{1 - P(X \leq \mathrm{Me})} \\ &= \frac{P(X < 115) - 0.5}{0.5}. \end{aligned}$$

Exercícios Resolvidos

Exercício 48., página 336

48.b) via `> pnorm(q,mean,sd)`

```
> (pnorm(115,100,15)-0.5)/0.5
[1] 0.6826895
```

48.b) via `> qnorm(q,mean,sd)` **&** `> pnorm(q,mean,sd)`

```
> Me<-qnorm(0.5,100,15)
> Me
[1] 100
> pMe<-pnorm(Me,100,15)
> p115<-pnorm(115,100,15)
> (p115-pMe)/(1-pMe)
[1] 0.6826895
```

Exercícios Resolvidos

Exercício 48., página 336

Variável aleatória em 48.c)

Corresponde uma v.a. do tipo binomial a $Y \sim B(n, \theta)$, sendo $n = 20$ o número de provas de Bernoulli ('20 pessoas') e $\theta = P(X > q_{0.75})$.

Obs: De **'Sucesso'='nota superior ao terceiro quartil'** **(**$Q3 = q_{0.75}$**)**, retira-se que $\theta = 1 - P(X \leq q_{0.75}) = 0.25$.

Probabilidade a calcular em 48.c)

Sendo $n = 20$, a probabilidade pretendida $P\left(Y \geq \frac{n}{2}\right) = P(Y \geq 10)$ pode ser determinada a partir de uma das seguintes fórmulas:

$P(Y \geq 10) =$	**Comando R**
$= P(Y > 9)$	$>$ `pbinom(9, 20, 0.25, lower.tail = FALSE)`
$= 1 - P(Y \leq 9)$	$>$ `1 - pbinom(9, 20, 0.25)`

Exercícios Resolvidos

Exercício 48., página 336

Variável aleatória em 48.c)

Corresponde uma v.a. do tipo binomial a $Y \sim B(n, \theta)$, sendo $n = 20$ o número de provas de Bernoulli ('20 pessoas') e $\theta = P(X > q_{0.75})$.

Obs:

'Sucesso'='nota superior ao terceiro quartil' $\left(Q3 = q_{0.75}\right)$ **.**

Probabilidade a calcular em 48.c)

Sendo $n = 20$, a probabilidade pretendida – pelo menos metade atinja nota superior a $Q3$ – pode ser determinada de uma das seguintes igualdades

$$\begin{aligned} P\left(Y \geq \frac{n}{2}\right) &= P(Y \geq 10) \\ &= 1 - P(Y \leq 9). \end{aligned}$$

Simplificação da probabilidade a calcular em 48.d)

Sendo $X \sim N(100,(15)^2)$ ($\mu = 100$ & $\sigma = 15$) $P((X-100)^2 \leq 144)$ pode ser reescrita, em termos de $P\left(\frac{X-100}{15} \leq z\right) = \Phi(z)$, da seguinte forma:

$$\begin{aligned} P((X-100)^2 \leq 144) &= \overbrace{P(-12 \leq X - 100 \leq 12)}^{=P(88 \leq X \leq 112)} \\ &= P\left(-\frac{12}{15} \leq \frac{X-100}{15} \leq \frac{12}{15}\right) \\ &= \Phi\left(\frac{12}{15}\right) - \Phi\left(-\frac{12}{15}\right) \\ &= 2(\Phi(0.8)-1) \quad \text{[Porquê?]} \end{aligned}$$

Exercícios Resolvidos

Exercício 48., página 336

48.d) via comando `>pnorm(q)`

```
> 2*pnorm(0.8)-1
[1] 0.5762892
> pnorm(12/15)-pnorm(-12/15)
[1] 0.5762892
```

48.d) via comando `>pnorm(q,mean,sd)`

```
> pnorm(112,100,15)-pnorm(88,100,15)
[1] 0.5762892
```

48.d) via comando `>pchisq(q,df)`

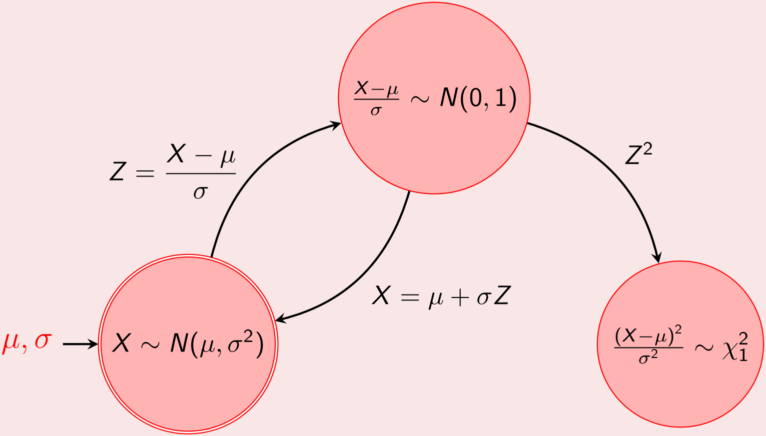
```
> pchisq(144/225,1)
[1] 0.5762892
```

Iremos falar mais adiante deste comando quando falarmos da distribuição qui-quadrado. Mais precisamente, da estandartização esquematizada na página seguinte:

Distribuição Normal $N(\mu, \sigma^2)$

Relação com a distribuição normal estandartizada, $N(0,1)$, e com a distribuição qui-quadrado [com 1 grau de liberdade]

$N(\mu, \sigma^2) \longrightarrow N(0,1) \longrightarrow \chi^2_1$



Exercícios Resolvidos

Exercício 49, páginas 336 – 337

Distribuição Exponencial de média 20

$$\frac{1}{\lambda} = E(X) = 20 \longrightarrow X \sim Ex\left(\frac{1}{20}\right) \longrightarrow F(x) = 1 - e^{-\frac{x}{20}} \; (x \geq 0)$$

49.a) `> pexp(q,rate)`

$$P(10 < Z \leq 30) =$$
$$= F(30) - F(10)$$
$$= e^{-\frac{1}{2}} - e^{-\frac{3}{2}}$$

[comando R]

```
> pexp(30, 1/20) - pexp(10, 1/20)
> exp(-1/2) - exp(-3/2)
```

49.b) 1-F(30) via `> pexp(q,rate,lower.tail=FALSE)`

```
> pexp(30,1/20,lower.tail = FALSE)
[1] 0.2231302
> exp(-3/2)
[1] 0.2231302
```

Exercícios Resolvidos

Exercício 49, páginas 336 – 337

49.c) via definição de probabilidade condicional

$$P(X > 40 | X > 10) = \frac{P(X > 40 \wedge X > 10)}{P(X > 10)} = \frac{P(X > 40)}{P(X > 10)}$$

49.c)> pexp(q,rate,lower.tail=FALSE)

```
> pexp(40,1/20,lower.tail =
FALSE)/pexp(10,1/20,lower.tail = FALSE)
[1] 0.2231302
> exp(-3/2)
[1] 0.2231302
```

49.c) via propriedade $P(X > x + h | X > x) = P(X > h)$

Tomando $x + h = 40$ e $x = 10$, segue que $h = 30 (= 40 - 10)$. Logo

$$P(X > 40 | X > 10) = P(X > 30)$$ **[solução 49.b)]**

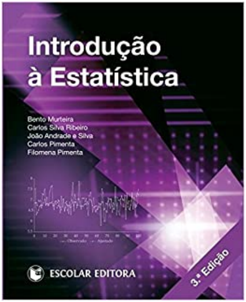
Conteúdos

Modelos Probabilísticos Usuais

Exercícios Propostos e Recomendados

Soluções Exercícios

- As soluções dos exercícios resolvidos e recomendados nos slides (**Extra-Aula**) encontram-se no final do livro [Murteira et al. (2015)].
- No caso dos exercícios associados ao tema **Modelos probabilísticos usuais**, as soluções destes encontram-se em [Murteira et al. (2015), pp. 782–785].



Murteira, B., C. Silva Ribeiro, J. Andrade e Silva, C. Pimenta, F. Pimenta (2015), **Introdução à Estatística, 3ª Edição**', Escolar Editora.

Conteúdos

Teste os seus conhecimentos

Distribuição de Poisson

Complementar ao Exercício 19.c), p 329

Seja $X \sim Po(3.2)$, e $f(x)$ a respectiva função de probabilidade.

1. Diga, justificando, se $f(3.2) = f(32.1)$.
2. Diga, justificando, para que valor de $x \in \mathbb{N}$, se tem $P(X \leq x) = P(X \geq 3.2)$

Complementar do Exercício 37.b), p. 333

Seja $M(s) = e^{e^s - 1}$ a função geradora de momentos de $X \sim Po(1)$, e $M_Y(s) = M(4s)$ a função geradora de momentos de $Y = 4X$.

1. Determine $E(Y)$ e $V(Y)$.
2. Diga, justificando, se $Y \sim Po(\lambda)$ para algum $\lambda > 0$.

Teste os seus conhecimentos

Distribuição Uniforme Contínua

Complementar do Exercício 35., p. 333

Mostre que se $X \sim U(a,b)$, então os quantis de ordem α são dados por $q_\alpha = a + (b-a)\alpha$.

Valores da média e variância conhecidos

Seja $Y \sim U(a,b)$ tal que $E(Y) = 58.5$ e $V(Y) = 168.75$.

1. Determine o valor dos parâmetros a & b de $U(a,b)$, e a função distribuição de probabilidade.
2. Determine o valor da probabilidade condicional $P(Y > 64 | Y \leq \text{Me})$, onde Me denota a mediana de $X \sim U(a,b)$.