4. Variáveis aleatórias especiais

4.1. Variáveis aleatórias discretas

4.1.1. Variável aleatória com distribuição uniforme discreta

- X v.a. discreta com $D = \{x_1, x_2, \dots, x_n\}$. Diz-se que X segue uma **distribuição uniforme** nos n pontos x_j $(j = 1,2,\dots,n)$ se e só se

$$f(x_j) = \frac{1}{n},\ j = 1,2,\dots,n.$$

- Cada valor x_j tem a mesma probabilidade.
- Momentos:

$$E(X) = \mu = \frac{1}{n}\sum_{j=1}^{n} x_j \qquad \mathrm{Var}(X) = \frac{1}{n}\sum_{j=1}^{n}(x_j - \mu)^2 = \frac{1}{n}\sum_{j=1}^{n} x_j^2 - \mu^2$$

$$E(X^k) = \frac{1}{n}\sum_{j=1}^{n} x_j^k$$

Dois casos particulares importantes:

1. $D = \{1, 2, ..., n\}$ Exemplo: Resultado do lançamento de um dado perfeito $(n = 6)$.

$$E(X) = \frac{n+1}{2}; \quad E(X^2) = \frac{(n+1)(2n+1)}{6}; \quad \mathrm{Var}(X) = \frac{n^2-1}{12};$$

2. $D = \{0, 1, 2, ..., m\}$ Exemplo: Escolha aleatória de um dos 10 dígitos $(m = 9)$.

$$E(X) = m/2; \quad E(X^2) = m(2m+1)/6; \quad \mathrm{Var}(X) = m(m+2)/12,$$

4.1.2. Variável aleatória com distribuição de Bernoulli

Prova de Bernoulli: Experiência aleatória em que se observa a realização ou não de determinado acontecimento *A;*

A realização de *A* designa-se por "sucesso" e a sua não realização por "insucesso". Define-se $P(A)=\theta$, $0<\theta<1$.

- Seja *X* a variável aleatória que caracteriza a experiência descrita. Assim,
  - Se $X=1$, o acontecimento *A* ocorre – verifica-se um "sucesso";
  - Se $X=0$, o acontecimento *A* não ocorre – verifica-se um "insucesso".

A função probabilidade de *X* escreve-se,

$$f(x\,|\,\theta)=\begin{cases}1-\theta & (x=0)\\ \theta & (x=1)\end{cases}$$

ou, de forma mais condensada,

$$f(x\,|\,\theta)=\theta^{x}(1-\theta)^{1-x},\quad x=0,1\quad (0<\theta<1).$$

- Se a v.a. X tem função probabilidade dada por

$$f(x\,|\,\theta)=\theta^{x}(1-\theta)^{1-x},\ x=0,1 \qquad (0<\theta<1)$$

diz-se que X tem **distribuição de Bernoulli**. Simbolicamente, $X \sim B(1;\theta)$.

- Note-se que se definiu uma *família de distribuições*. A cada valor de θ corresponde uma distribuição concreta.

- Facilmente se verifica que:

$$E(X)=\theta\,;\quad E(X^2)=\theta\,;\quad \mathrm{Var}(X)=\theta(1-\theta);\quad \sigma=\sqrt{\theta(1-\theta)}\quad \gamma_1=(1-2\theta)/\sigma\,.$$

4.1.3 Variável aleatória com distribuição binomial

- Qual a probabilidade para que, em ***n*** **provas de Bernoulli independentes**, se obtenham x sucessos $(x = 0,1,\ldots,n)$ seja qual for a ordem em que estes são obtidos?

Considere-se a probabilidade de x sucessos seguidos de $n - x$ insucessos,

$\underbrace{AA\ldots A}_{x}\,\underbrace{\bar{A}\bar{A}\ldots\bar{A}}_{n-x}$ e considerem-se todas as ordens possíveis;

Obtém-se $\binom{n}{x}\theta^x(1-\theta)^{n-x}$ (esquema binomial visto no capítulo 1)

- **Exemplo 4.1 –** Qual a probabilidade para, em cinco lançamentos de um dado "perfeito", se obter duas vezes a «sena»?

- Se a v.a. X tem função probabilidade

$$f(x|\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x}, \quad x = 0,1,2,\ldots,n \quad (0 < \theta < 1),$$

diz-se que tem **distribuição binomial**. Simbolicamente, $X \sim B(n;\theta)$.

  - Família de distribuições, indexadas pelos parâmetros n e θ;

- A distribuição de Bernoulli pode obter-se da distribuição binomial fazendo $n = 1$.
- Cálculo das probabilidades para distribuições binomiais:
    - Máquina calcular ou computador
    - Tabelas estatísticas

- Tabela 1 para valores de n de 1 a 20 e para alguns valores de θ (0.05, 0.10, 0.15, …, 0.50). Quando $\theta > 0.50$ a relação, $X \sim B(n;\theta) \Leftrightarrow Y = (n - X) \sim B(n;1-\theta)$, permite remeter todos os cálculos para as tabelas construídas para $0 < \theta < 0.50$.
- **Exemplo 4.2 –** Sabe-se que, com determinado tratamento administrado a pacientes em condições bem definidas, se alcançam 70% de curas para certa doença. Se o tratamento é aplicado a 20 pacientes em tais condições, qual é a probabilidade de: (a) obter 15 curas no máximo; (b) obter 12 ou mais curas; (c) obter um número de curas não inferior a 10 nem superior a 15.

- Momentos

$$E(X) = n\theta$$

$$\mathrm{Var}(X) = n\theta(1-\theta)$$

$$\sigma = \sqrt{n\theta(1-\theta)}$$

$$\gamma_1 = \frac{1-2\theta}{\sigma}$$

Binomial e Bernoulli

X_i $(i = 1,2, \dots, n)$: v.a. associada com a i-ésima prova de Bernoulli (X_i tem distribuição de Bernoulli com parâmetro θ).

Se adicionalmente as variáveis aleatórias X_i forem independentes, a variável $X = \sum_{i=1}^{n} X_i$ tem distribuição binomial de parâmetros n e θ

Teorema – Sejam Y_1 e Y_2 duas variáveis aleatórias independentes. Então,

$Y_1 \sim B(n_1; \theta), Y_2 \sim B(n_2; \theta) \Rightarrow Y = Y_1 + Y_2 \sim B(n, \theta)$,

onde $n = n_1 + n_2$.

4.1.4. Variável aleatória com distribuição de Poisson

Uma variável aleatória X tem distribuição de Poisson se e só se,

$$f(x|\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!}, \quad x = 0,1,2,\ldots, \quad (\lambda > 0).$$

Simbolicamente, $X \sim \mathrm{Po}(\lambda)$.

Função de distribuição

$$F(x|\lambda) = P(X \le x) = \sum_{i=0}^{x} \frac{e^{-\lambda}\lambda^{i}}{i!}$$

(se x é um inteiro)

Momentos: Média igual à variância

$E(X) = \lambda;$ $\quad$ $\mathrm{Var}(X) = \lambda;\ \sigma = \sqrt{\lambda};\ \gamma_1 = \lambda^{-1/2}$

- **Exemplo 4.3 –** Numa fábrica de têxteis existem numerosos teares do mesmo tipo. Concluiu-se que o número de teares que se avariam cada mês é uma variável aleatória $X \sim \mathrm{Po}(\lambda = 3)$.

  Calcule a probabilidade para que durante um mês se avariem sete ou mais teares:

  Determinar a capacidade mensal mínima disponível C da oficina de reparação, de modo a ser pelo menos 0.9 a probabilidade de não haver teares aguardando reparação.

Soma de Poisson independentes

- Teorema – Sejam X_1 e X_2 duas v.a. independentes. Então,

$X_1 \sim \mathrm{Po}(\lambda_1), \quad X_2 \sim \mathrm{Po}(\lambda_2) \Rightarrow X = X_1 + X_2 \sim \mathrm{Po}(\lambda)$ onde $\lambda = \lambda_1 + \lambda_2$.

Lei dos Acontecimentos raros → Binomial para Poisson.

Quando $\theta = \lambda/n \to 0$, mantendo-se fixo $n\theta = \lambda$, a binomial tende para a Poisson,

$$\lim_{n\to\infty} \binom{n}{x}\theta^x(1-\theta)^{n-x} = \lim_{n\to\infty} \binom{n}{x}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x} = \frac{e^{-n\theta}(n\theta)^x}{x!}.$$

A **regra prática** para utilizar esta "lei" baseia-se no pressuposto de que se tem um acontecimento raro e um número "elevado" de observações.

Assim, **não é aconselhável** fazer a aproximação quando:

- $0.1 < \theta < 0.9$ (quando $\theta \geq 0.9$, evidentemente que o acontecimento em causa não é "raro", mas sim o seu complementar)
- $n \leq 20$ (que são os valores de *n* considerados na tabela 1).

- **Exemplo 4.4** Sabendo que $\theta = 0.001$ é a probabilidade, *P*, de uma peça, produzida por certa máquina, ser defeituosa, qual a probabilidade de, num lote de 1000 peças haver mais de uma defeituosa?

Processo de Poisson e distribuição de Poisson

- A distribuição de Poisson está associada com um processo de contagem com o mesmo nome;
- Definição de um processo de Poisson – Suponha-se que se procede à contagem do número de eventos ocorridos ao longo do tempo. Tem-se um processo de Poisson com parâmetro (taxa média) $\lambda > 0$ quando se verificam as seguintes condições:
  - o número de eventos que ocorrem em dois intervalos disjuntos são independentes;
  - a probabilidade de ocorrer exatamente um evento em qualquer intervalo de amplitude Δt arbitrariamente pequena é aproximadamente $\lambda\Delta t$;
  - a probabilidade de ocorrerem dois ou mais eventos em qualquer intervalo de amplitude Δt arbitrariamente pequena é aproximadamente igual a zero.

Num **processo de Poisson**, os acontecimentos ocorrem a uma taxa média de λ por unidade de tempo. É possível demonstrar que o número de ocorrências num intervalo de amplitude t tem distribuição de Poisson de parâmetro λt

$$f(x|\lambda t) = \frac{e^{-\lambda t}\,(\lambda t)^x}{x!}, \qquad x = 0,1,2,\ldots \quad (\lambda > 0).$$

Exemplo: Considere que o número de pessoas que compram, pelo menos, uma peça de vestuário na loja FORADEMODA segue um processo de Poisson com taxa média de 3 por dia. Calcule a probabilidade de 10 ou mais pessoas adquirirem, pelo menos, uma peça de vestuário na loja FORADEMODA durante uma semana (5 dias).

4.2. Variáveis aleatórias contínuas

4.2.1. Variável aleatória com distribuição uniforme contínua

- Mesmo nome (porque a ideia é a mesma) mas diferente da uniforme discreta
- A v.a. X tem distribuição uniforme no intervalo (α, β), com $\alpha < \beta$, quando a função densidade é da forma,

$$f(x|\alpha,\beta) = \begin{cases} \frac{1}{\beta-\alpha} & (\alpha < x < \beta), \\ 0 & (\text{outros } x). \end{cases}$$

Simbolicamente, $X \sim U(\alpha, \beta)$.

- Função de distribuição:

$$F(x|\alpha,\beta) = \begin{cases} 0 & (x < \alpha), \\ \frac{x-\alpha}{\beta-\alpha} & (\alpha \leq x < \beta), \\ 1 & (x \geq \beta). \end{cases}$$

- Expressão geral dos momentos em relação à origem

$$E\left(X^k\right) = \frac{\beta^{k+1}-\alpha^{k+1}}{(k+1)(\beta-\alpha)}.$$

$$E(X) = \frac{\alpha+\beta}{2}, \quad \mathrm{Var}(X) = \frac{(\beta-\alpha)^2}{12}, \quad \gamma_1 = 0$$

Uniforme (0, 1)

O caso $\alpha = 0, \beta = 1$, isto é, $X \sim U(0,1)$, é o de **maior interesse**

$f(x) = \begin{cases} 1 & (0 < x < 1) \\ 0 & (\text{outros } x) \end{cases}$	$F(x) = \begin{cases} 0 & (x < 0) \\ x & (0 \le x < 1) \\ 1 & (x \ge 1) \end{cases}$
$E(X^k) = \frac{1}{(k+1)}$ logo $E(X) = \frac{1}{2}$, $\text{Var}(X) = \frac{1}{12}$, $\gamma_1 = 0$	

- **Teorema – Transformação uniformizante – resultado particularmente importante em problemas de simulação.**

  Este resultado mostra que, em certas condições, $Y = F_X(X) \sim U(0,1)$ e inversamente que se $Y \sim U(0,1)$ então $X = F_X^{-1}(Y) \sim F_X(x)$

Exemplo 4.5 – Suponha-se que a variável aleatória X tem função de distribuição é $F_X(x) = 1 - e^{-x}$. Determinar-se a função distribuição da variável aleatória $Y = F_X(X) = 1 - e^{-X}$.

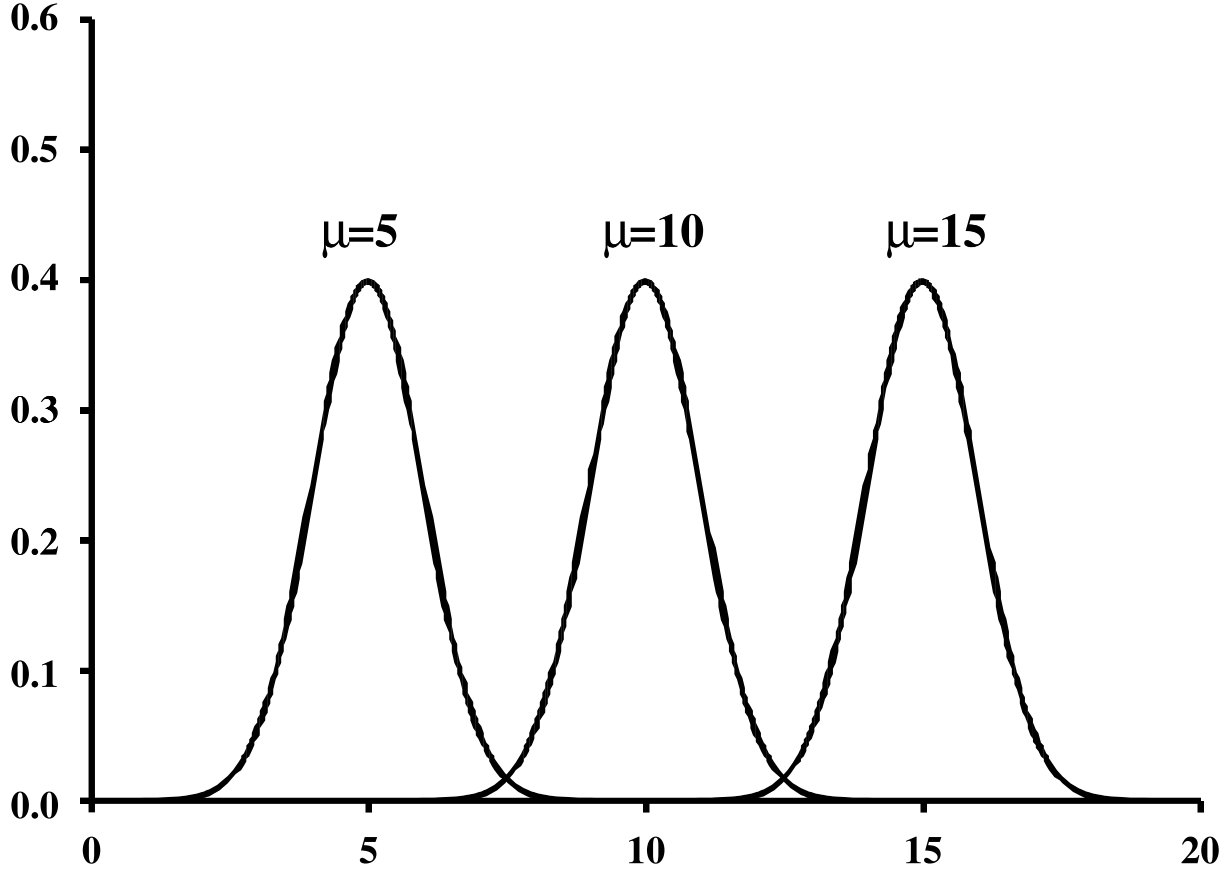
4.2.2. – Variável aleatória com distribuição normal

- Os parâmetros da distribuição normal representam-se por μ e σ^2 porque correspondem respetivamente, como se vai ver, à média e variância da variável aleatória *X*.
- A v.a. *X* tem distribuição normal com parâmetros μ e σ^2 quando a função densidade é da forma

$$f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}, -\infty < x < +\infty, \quad -\infty < \mu < +\infty, \quad 0 < \sigma^2 < +\infty,$$

Simbolicamente $X \sim N(\mu,\sigma^2)$.

- Função de distribuição: $F(x|\mu,\sigma^2) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{1}{2\sigma^2}(t-\mu)^2\right\} dt$, para o qual **não se conhece solução analítica**



Fig– Funções densidade da distribuição normal com a mesma variância ($\sigma^2 = 1$)

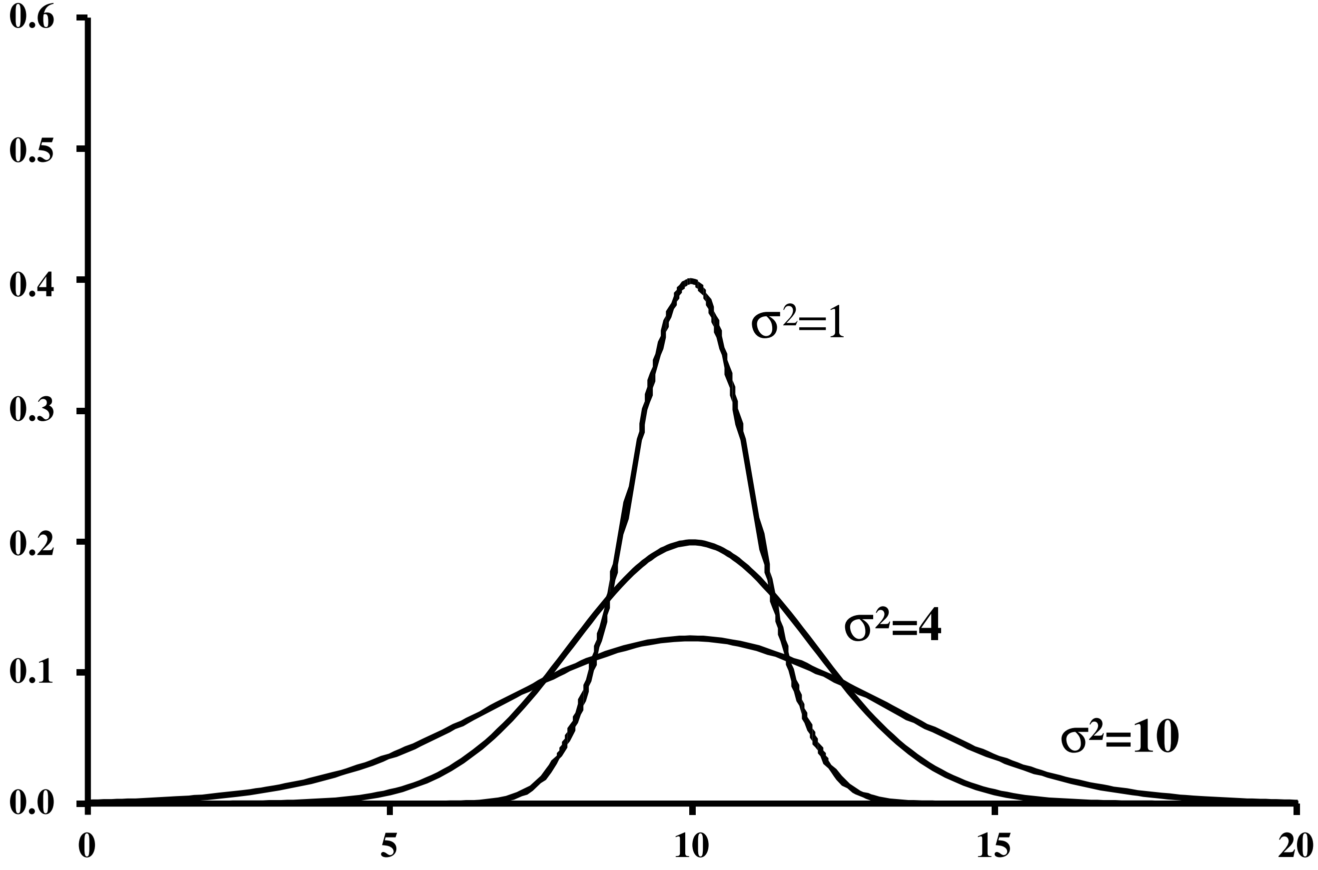


Fig. – Funções densidade da distribuição normal com a mesma média ($\mu = 10$)

Distribuição normal estandardizada - caso particular, com $\mu = 0$ e $\sigma^2 = 1$

- função densidade $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$
- função de distribuição $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt$

Facilmente se passa de uma normal (μ, σ^2) para uma normal $(0,1)$ já que, se $X \sim N(\mu, \sigma^2)$, a variável estandardizada,

$$Z = \frac{X - \mu}{\sigma} \quad \sim N(0,1)$$

Propriedades da distribuição normal

- A função densidade normal é simétrica em relação μ,

$$f(\mu - x) = f(\mu + x)$$

Em particular, tratando-se da distribuição normal estandardizada,

$$\phi(x) = \phi(-x), \quad \Phi(x) = 1 - \Phi(-x)$$ importante

Momentos

- $E(X) = \mu$, $\text{Var}(X) = \sigma^2$.
- Os momentos centrais de ordem impar são todos nulos (simetria)
- a expressão dos momentos centrais de ordem par é $\mu_{2r} = \frac{(2r)!\sigma^{2r}}{2^r r!}$, $r = 1,2,\cdots$
- Logo $\gamma_1 = 0$, $\mu_4 = 3\sigma^4$ e portanto $\gamma_2 = 3$

- É habitual classificar as distribuições simétricas comparando-as com a normal:
  - Distribuições *leptokurtica* ($\gamma_2 > 3$) → caudas mais "espessas" (com zona central mais "pontiaguda") que a distribuição normal.
  - Distribuições *platikurtica* ($\gamma_2 < 3$)→ caudas mais "finas" (com zona central mais "achatada") que a distribuição normal.
  - Distribuição *mesokurtica* ($\gamma_2 = 3$)

Tabelas da distribuição normal estandardizada (padronizada):

A tabela 4 refere-se à função $\Phi(z)$

A tabela 5 diz respeito à função inversa $\Phi^{-1}(z)$, permitindo determinar, para certos valores de $\Phi(z)$, a respectiva abcissa z.

Exemplo:

- $X \sim N(0,1)$ Calcular $P(X < 2.53)$, $P(X < -2.45)$, $P(-1 < X < 2.03)$
- $X \sim N(1,4)$ Calcular $P(X < 2.53)$, $P(X < -2.45)$, $P(-1 < X < 2.03)$

Exemplo 4.6 – A resistência à compressão de amostras de cimento de um certo tipo é uma variável aleatória que pode ser modelada por uma distribuição normal com média 6000 kg/cm^2 e desvio padrão 100 kg/cm^2.

Calcule

a) a probabilidade para que uma amostra de cimento tenha resistência superior a 6150 kg/cm^2.

b) a probabilidade de que uma amostra de cimento tenha resistência entre 5900 kg/cm^2 e 5950 kg/cm^2.

c) a resistência R que é excedida por 90% das amostras de cimento.

Teorema – Aditividade da normal

- Se as variáveis aleatórias $X_i,\ i = 1,2,\ldots,k$, são independentes então,

$$X_i \sim N(\mu_i, \sigma_i^2) \Rightarrow Y = \sum_{i=1}^{k} \alpha_i X_i \sim N(\mu_Y, \sigma_Y^2),$$

onde,

$$\mu_Y = \sum_{i=1}^{k} \alpha_i \mu_i \text{ e } \sigma_Y^2 = \sum_{i=1}^{k} \alpha_i^2 \sigma_i^2.$$

Corolário 1 – Se as variáveis aleatórias $X_i,\ i = 1,2,\ldots,k$, são independentes e identicamente distribuídas então

$$X_i \sim N(\mu, \sigma^2) \Rightarrow Y = \sum_{i=1}^{k} X_i \sim N(k\mu, k\sigma^2)$$

Corolário 2 – Se as variáveis aleatórias $X_i,\ i = 1,2,\ldots,k$, são independentes e identicamente distribuídas

$$X_i \sim N(\mu, \sigma^2) \Rightarrow \bar{X} = \frac{1}{k}\sum_{i=1}^{k} X_i \sim N\left(\mu, \frac{\sigma^2}{k}\right).$$

4.2.3. – Variável aleatória com distribuição exponencial

- Diz-se que a v.a. X tem distribuição exponencial de parâmetro λ quando a sua função densidade é da forma,

$$f(x|\lambda) = \begin{cases} 0\ , & x \leq 0, \\ \lambda e^{-\lambda x}, & x > 0. \end{cases}$$

Simbolicamente, $X \sim \mathrm{Ex}(\lambda)$.

- Função de distribuição: $F(x) = \begin{cases} 0 & (x \leq 0), \\ 1 - e^{-\lambda x} & (x > 0), \end{cases}$

- $E(X) = \frac{1}{\lambda}$, $\mathrm{Var}(X) = \frac{1}{\lambda^2}$, $\mathrm{CV} = 1$, $\gamma_1 = 2$, $\gamma_2 = 9$.

- Moda: **Não existe** (a função densidade é decrescente com x e o domínio é aberto).

- Mediana: $\mu_e = ln(2)/\lambda$. Note-se que $\mu_e < \mu$

Duas <u>propriedades</u>:

- "Falta de memória": $X \sim \text{Ex}(\lambda) \Rightarrow P(X > x + h | X > x) = P(X > h)$.
- $X_1, X_2, \ldots, X_k$ v.a. **independentes** então, $X_i \sim \text{Ex}(\lambda) \Rightarrow Y = \min_i X_i \sim \text{Ex}(k\lambda)$.

Exemplo 4.7 Suponha que o tempo de vida útil de um dado componente é uma v.a. *X* com distribuição exponencial com média igual a 600 horas.

- Qual é a probabilidade de o componente durar mais de 700 horas?
- Sabendo que o componente já durou mais de 400 horas, qual é a probabilidade de durar pelo menos ainda mais 700 horas?
- Num conjunto de 10 componentes a funcionar de forma independente, todos com tempo de vida útil com a mesma distribuição de X, qual a probabilidade da menor vida útil ser inferior a 100 horas?

- Propriedade: Se uma sucessão de eventos constitui um processo de Poisson de intensidade λ e se se inicia a contagem no instante 0, o **tempo de espera** pela chegada da primeira ocorrência é uma variável aleatória, com distribuição exponencial de parâmetro λ.

Exemplo 4.8 – A chegada de clientes a uma loja segue um processo de Poisson em que o ritmo médio de afluência é de 20 clientes por hora. Após abrir a loja, qual é a probabilidade de o comerciante ter de esperar mais de 5 minutos pela chegada do primeiro cliente?

4.2.4– Variáveis aleatórias com distribuição gama e distribuição do qui-quadrado

- Uma variável aleatória X com função densidade dada por,

$$f(x|\alpha,\lambda) = \frac{\lambda^{\alpha} e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)}, x > 0, \quad \alpha > 0, \lambda > 0,$$

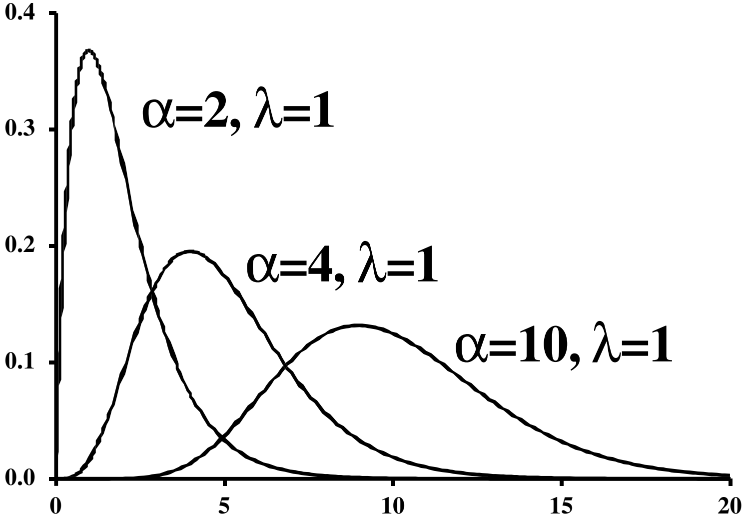
diz-se ter **distribuição gama** de parâmetros α e λ. Simbolicamente, $X \sim G(\alpha,\lambda)$.

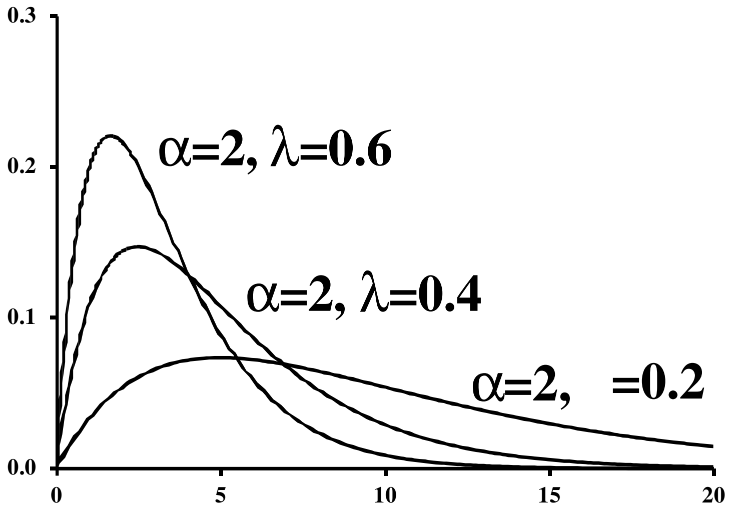
- **Função gama** ou fatorial generalizado:

$$\Gamma(\alpha) = \int_0^{+\infty} e^{-x} x^{\alpha-1} dx \qquad (\alpha > 0)$$

Propriedades:

- $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1), \quad \alpha > 1.$
- Se n inteiro, $\Gamma(n) = (n-1)!$
- $\Gamma(1) = 1$ e $\Gamma(1/2) = \sqrt{\pi}$





Casos Particulares:

- Quando $\alpha = 1$, temos a distribuição exponencial.
- Se $\alpha = n/2$ e $\lambda = 1/2$ temos a dist. **Qui-quadrado**, i.e. $\chi^2(n) \equiv G\left(\frac{n}{2};\frac{1}{2}\right)$.

Momentos

- $E(X^k) = \frac{\alpha(\alpha+1)\ldots(\alpha+k-1)}{\lambda^k}$

$$E(X) = \mu = \frac{\alpha}{\lambda}, \quad \mathrm{Var}(X) = \sigma^2 = \frac{\alpha}{\lambda^2}, \quad \mathrm{CV} = \frac{1}{\sqrt{\alpha}}\ , \ \gamma_1 = \frac{2}{\sqrt{\alpha}}, \quad \gamma_2 = 3 + \frac{6}{\alpha}.$$

Parâmetros de ordem

- Moda: a distribuição gama não tem moda quando $\alpha \leq 1$. Quando $\alpha > 1$,

$$\mathrm{moda}(X) = \mu_* = (\alpha - 1)/\lambda.$$

- Mediana: existe sempre, embora não se consiga explicitar analiticamente a sua expressão.
- Quando existem as três medidas de localização ($\alpha > 1$), verifica-se a seguinte relação: $\mu_* < \mu_e < \mu$.

Aditividade da gama (obriga à independência e ao mesmo parâmetro λ)

- Sejam X_1 e X_2 duas variáveis aleatórias **independentes**. Então,

$$X_1 \sim G(\alpha_1;\lambda),\quad X_2 \sim G(\alpha_2;\lambda) \Rightarrow X = X_1 + X_2 \sim G(\alpha;\lambda),\quad \alpha = \alpha_1 + \alpha_2.$$

- Este resultado pode ser generalizado para k v.a. **independentes**

- A distribuição gama com $\alpha = n$ inteiro pode ser interpretada como a soma de α exponenciais independentes e portanto, em termos do processo de Poisson podemos interpretar a Gama como o tempo de espera a distribuição do tempo de espera pelo n -ésimo evento.

Cálculo de probabilidades

- Não existe, em geral, solução analítica do integral correspondente à função de distribuição logo recorre-se ao computador
- A distribuição gama não se encontra tabelada. Para alguns casos particulares recorre-se à distribuição do qui-quadrado

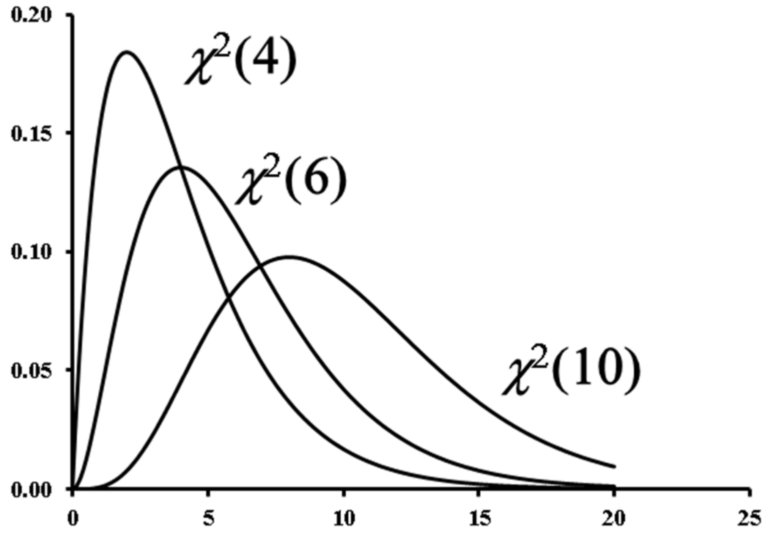
Distribuição do qui-quadrado

Diz-se que a v.a. X tem distribuição do qui-quadrado com n graus de liberdade (n inteiro positivo), simbolicamente $X \sim \chi^2(n)$, quando a respetiva função densidade é da forma,

$$f(x|n) = \frac{e^{-x/2} x^{n/2-1}}{2^{n/2}\Gamma(n/2)}, x > 0, n > 0$$

A distribuição do qui-quadrado é uma distribuição gama com $\alpha = n/2$ e $\lambda = 1/2$,

$X \sim \chi^2(n) \Leftrightarrow X \sim G\left(\frac{n}{2}; \frac{1}{2}\right)$.



Resultado importante

$$X \sim G(n;\lambda) \Leftrightarrow Y = 2\lambda X \sim \chi^2(2n)$$

Como calcular probabilidades?

- Máquina de calcular ou computador
- Tabelas:

$X \sim \chi^2(n)$ A tabela permite encontrar o valor $\chi^2_{n,\varepsilon}$ tal que, $P(X > \chi^2_{n,\varepsilon}) = \varepsilon$, para alguns valores de ε e de n.

Para valores que não estão contemplados na tabela 6 e $n > 100$, pode utilizar-se

$$X \sim \chi^2(n) \Rightarrow \sqrt{2X} - \sqrt{2n-1} \overset{a}{\sim} N(0,1)$$

- **Exemplo 4.8** Seja $X \sim G(2; 0.008). Calcular\ P(X > 500)$

Momentos

$$E(X) = n, \quad \mathrm{Var}(X) = 2n, \quad \mathrm{CV} = \sqrt{\frac{2}{n}}, \quad \gamma_1 = \sqrt{\frac{8}{n}}, \quad \gamma_2 = 3 + \frac{12}{n},$$

Aditividade da qui-quadrado (obriga à **independência**)

X_1 e X_2 duas v.a. independentes. Então,

$$X_1 \sim \chi^2(n_1), \quad X_2 \sim \chi^2(n_2) \Rightarrow X = X_1 + X_2 \sim \chi^2(n), n = n_1 + n_2.$$

- Pode generalizar-se par k v.a. independentes

Relação entre a normal e a qui-quadrado

- Caso 1: $X \sim N(0,1) \Rightarrow Y = X^2 \sim \chi^2(1)$.
- Caso 2: $X_i \sim N(0,1) \Rightarrow \sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$ se $X_i,\ i = 1,2,\ldots,n$, independentes.
- Caso 3: $X_i \sim N(\mu_i, \sigma_i^2) \Rightarrow \sum_{i=1}^{n} \left(\frac{X_i - \mu_i}{\sigma_i}\right)^2 \sim \chi^2(n)$ se $X_i,\ i = 1,2,\ldots,n$, independentes.

4.3. Teorema do limite central

- O T.L.C é um dos **resultados mais importantes** da teoria da probabilidade.

Teorema [Teorema do limite central (Lindberg-Levy)] – Dada a sucessão de variáveis aleatórias *iid*, $X_1, X_2, \ldots, X_n, \ldots$, com média μ e variância σ^2, então, quando $n \to +\infty$, a função de distribuição da variável aleatória,

$$Z_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}},$$

tende para uma função de distribuição $N(0,1)$, ou seja, a distribuição assintótica de Z_n é a $N(0,1)$. Simbolicamente, $Z_n \overset{a}{\sim} N(0,1)$.

Interpretação prática do TLC

A conclusão do teorema anterior pode exprimir-se na forma alternativa,

- $\lim\limits_{n\to+\infty} P(Z_n \le x) = \lim\limits_{n\to+\infty} P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \le x\right) = \Phi(x),$ ou

se n “grande”:

- $P(Z_n \le x) = P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \le x\right) \approx \Phi(x)$ (n grande).

Basta exigir a existência de μ e σ^2. Distribuição de X_i pode ser discreta ou contínua.

O que é *n* grande? Depende... da distribuição de X_i.

- Distribuições simétricas e unimodais tornam a convergência mais rápida e melhoram a aproximação;
- Aconselhável no mínimo $n \geq 30$, salvo nalguns casos muito especiais (c/ aplicação prática). Para certas distribuições ter-se-ão de observar valores de *n* muito mais elevados.
- Alguns cuidados quando a distribuição de X_i é discreta: Correção de continuidade.

Exemplo 4.9 – A procura diária a satisfazer (unidade: 100 Kg) é uma v.a. X com média 40 e variância 25. Sendo a produção anual planeada é de 11500, calcular a probabilidade de haver procura anual excedentária (ano: 289 dias úteis).

Exemplo 4.10 – Do exemplo anterior pretende determinar-se a produção Q que deve ser planeada de modo a cobrir a procura anual com probabilidade 0.99.

T. Limite Central: Aplicação a aproximações para distribuições discretas

Corolário (T. de De Moivre-Laplace) – Dada a sucessão de variáveis aleatórias iid, $X_1, X_2, \ldots, X_n, \ldots$, com distribuição de Bernoulli de média $E(X_i) = \theta$ e, portanto, $\mathrm{Var}(X_i) = \theta(1-\theta)$, tem-se,

$$\frac{\sum_{i=1}^{n} X_i - n\theta}{\sqrt{n\theta(1-\theta)}} \overset{a}{\sim} N(0,1).$$

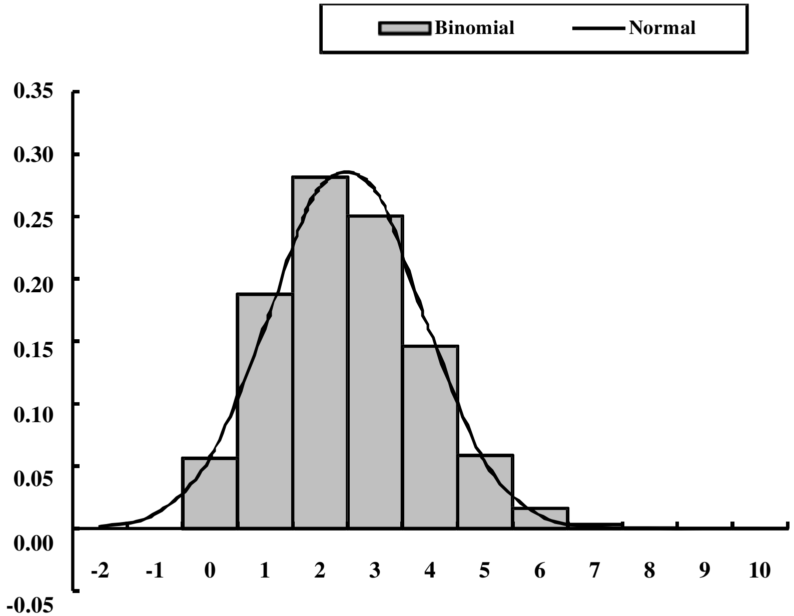


Fig. – Aproximação da binomial pela normal: $n = 10$, $\theta = 0.25$.

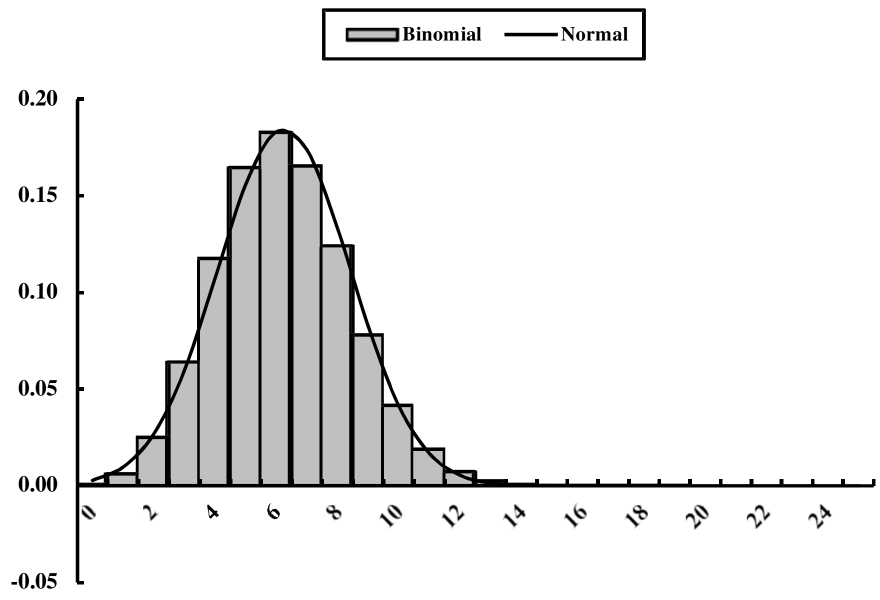


Fig.– Aproximação da binomial pela normal: $n = 25, \theta = 0.25$.

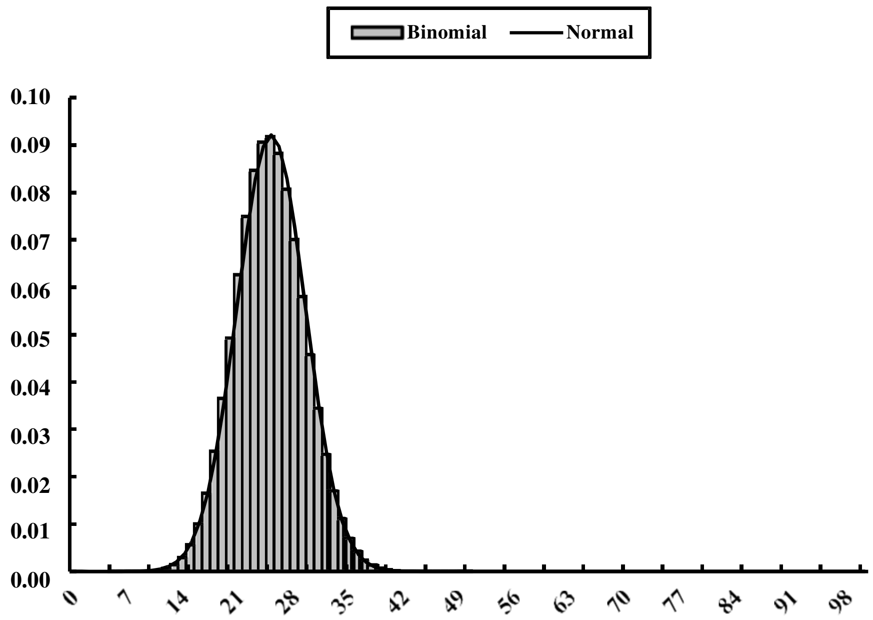


Fig. – Aproximação da binomial pela normal: $n = 100$, $\theta = 0.25$.

Nota: $X = \sum_{i=1}^{n} X_i \sim B(n;\theta)$. Quando n é grande, utilizar o corolário. Mas...

com Correção de Continuidade.

$X \sim B(n;\theta)$ (n grande) ,$P(a \leq X \leq b) =?$, a e b inteiros, $0 \leq a < b \leq n$.

- Resposta exata: $P(a \leq X \leq b) = \sum_{x=a}^{b} \binom{n}{x} \theta^x (1-\theta)^{n-x}$
- Resposta aproximada: $P(a \leq X \leq b) \approx \Phi\left(\frac{b-n\theta}{\sqrt{n\theta(1-\theta)}}\right) - \Phi\left(\frac{a-n\theta}{\sqrt{n\theta(1-\theta)}}\right)$,

pois $\frac{X-n\theta}{\sqrt{n\theta(1-\theta)}} \overset{a}{\sim} N(0,1)$,

Correcção de continuidade

$$P(a \leq X \leq b) \approx \Phi\left(\frac{b+\frac{1}{2}-n\theta}{\sqrt{n\theta(1-\theta)}}\right) - \Phi\left(\frac{a-\frac{1}{2}-n\theta}{\sqrt{n\theta(1-\theta)}}\right).$$

A **correção de continuidade aplica-se sempre quando se aproxima uma distribuição discreta por uma distribuição contínua**.

Cálculo de probabilidades com a binomial (regra prática):

- Sempre que possível utilizar diretamente a distribuição binomial.
- Quando necessário o cálculo aproximado das probabilidades deve atender aos seguintes casos:
  - Se $0.1 < \theta < 0.9 \rightarrow$ (ou $n\theta \geq 5$) aproximar pela normal, utilizando a correção de continuidade.
  - Se $\theta \leq 0.1 \rightarrow$ aproximar pela Poisson.
  - Se $\theta \geq 0.9 \rightarrow$ aproximar pela Poisson, considerando o respectivo acontecimento complementar.

Exemplo 4.11 – Seja $X \sim B(200; 0.5)$. Calcular $P(95 \leq X \leq 105)$

Corolário – Se X é uma variável com distribuição de Poisson, $X \sim \mathrm{Po}(\lambda)$, então,

$$\lim_{\lambda \to +\infty} P\left(\frac{X-\lambda}{\sqrt{\lambda}} \leq x\right) = \Phi(x).$$

Efectuar a correcção de continuidade.

Exemplo 4.12 – Suponha-se que $X \sim \mathrm{Po}(20)$. Calcular $P(16 < X < 22)$