FORMULÁRIO PROBABILIDADE E ESTATÍSTICA

Navegação rápida por fórmulas

Índice geral: Página 2
Índice de fórmulas: Página 3

Clique-se diretamente na fórmula desejada para navegar até lá

Autor: Vicente Duarte

Data: 7 de julho de 2025

Índice Geral

Conteúdo

Índice de Fórmulas

Clique na fórmula desejada para navegar diretamente:

1 Conceitos básicos de probabilidade

1.1 Experiência aleatória. Espaço de resultados e acontecimentos

1.2 Noção de probabilidade. Probabilidade condicionada e lei da probabilidade total

Acontecimentos mutuamente exclusivos

Dois acontecimentos são mutuamente exclusivos se não podem ocorrer simultaneamente:

$$A \cap B = \emptyset \quad \Rightarrow \quad P(A \cap B) = 0 \qquad \text{(Fórmula 1)}$$

Probabilidade da união de três acontecimentos

$$\begin{aligned} P(A \cup B \cup C) = {} & P(A) + P(B) + P(C) \\ & - P(A \cap B) - P(A \cap C) - P(B \cap C) \\ & + P(A \cap B \cap C) \end{aligned} \qquad \text{(Fórmula 2)}$$

Fórmula da inclusão-exclusão para n acontecimentos

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n+1} P(A_1 \cap \cdots \cap A_n)$$

(Fórmula 3)

1.3 Teorema de Bayes

O Teorema de Bayes permite calcular a probabilidade *a posteriori* de um acontecimento A_i dado que ocorreu um acontecimento B.

Teorema de Bayes

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{\sum_{j=1}^{n} P(B \mid A_j)P(A_j)} \qquad \text{(Fórmula 4)}$$

Onde:

- $A_1, \ldots, A_n$ formam uma partição do espaço amostral
- B é um acontecimento tal que $P(B) > 0$
- $P(A_i)$ é a probabilidade *a priori* de A_i
- $P(B \mid A_i)$ é a verosimilhança de B dado A_i
- $P(A_i \mid B)$ é a probabilidade *a posteriori* de A_i dado B

1.4 Acontecimentos independentes

Independência de acontecimentos

Dois acontecimentos A e B são independentes se e só se:

$$P(A \cap B) = P(A) \cdot P(B) \qquad \text{(Fórmula 5)}$$

2 Variáveis aleatórias discretas e contínuas

2.1 Definição de variável aleatória. Função de distribuição. Função de massa de probabilidade e função de densidade de probabilidade

2.2 Valor esperado, moda, variância e quantis

Valor esperado para variáveis discretas

$$E(X) = \sum_x x\,P(X = x) \qquad \text{(Fórmula 6)}$$

Valor esperado para variáveis contínuas

$$E(X) = \int_{-\infty}^{+\infty} x\,f_X(x)\,dx \qquad \text{(Fórmula 7)}$$

Variância

$$\mathrm{Var}(X) = E[(X - E(X))^2] = E(X^2) - [E(X)]^2 \qquad \text{(Fórmula 8)}$$

Variância amostral

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - n\,\overline{x}^2}{n-1} \qquad \text{(Fórmula 8a)}$$

Propriedades da esperança e variância

Esperança:

$$E(aX + bY) = aE(X) + bE(Y) \qquad \text{(Fórmula 9)}$$

Variância:

$$\mathrm{Var}(aX + bY) = a^2\,\mathrm{Var}(X) + b^2\,\mathrm{Var}(Y) + 2ab\,\mathrm{Cov}(X, Y) \qquad \text{(Fórmula 10)}$$

Quantis

Mediana: valor m tal que $P(X \leq m) \geq 0.5$ e $P(X \geq m) \geq 0.5$
Quantil de ordem q $(0 < q < 1)$:

$$x_q = F_X^{-1}(q) \quad \text{onde } F_X(x_q) = q \qquad \text{(Fórmula 11)}$$

2.3 Distribuições de probabilidade mais utilizadas na modelação de dados

Propriedade de falta de memória - Distribuição geométrica

Para uma sequência de ensaios de Bernoulli:

$$P(X > m + n \mid X > m) = P(X > n) \qquad \text{(Fórmula 12)}$$

Interpretação: A probabilidade de esperar mais n tentativas, dado que já se esperou m, é independente do número de tentativas já realizadas.

Propriedade de falta de memória - Distribuição exponencial

Se $T \sim \text{Exp}(\lambda)$ representa o tempo até à ocorrência de um evento:

$$P(T > t + s \mid T > s) = P(T > t) \qquad \text{(Fórmula 13)}$$

Nota: A distribuição exponencial é a única distribuição contínua com propriedade de falta de memória.

Soma de variáveis binomiais independentes

Se $X_i \sim \text{Bin}(n_i, p)$ são independentes, então:

$$X = \sum_{i=1}^{k} X_i \sim \text{Bin}\left(\sum_{i=1}^{k} n_i,\ p\right) \qquad \text{(Fórmula 14)}$$

Condição: Todas as variáveis devem ter o mesmo parâmetro p.

Soma de variáveis de Poisson independentes

Se $X_i \sim \text{Poisson}(\lambda_i)$ são independentes, então:

$$X = \sum_{i=1}^{k} X_i \sim \text{Poisson}\left(\sum_{i=1}^{k} \lambda_i\right) \qquad \text{(Fórmula 15)}$$

Combinação linear de variáveis normais independentes

Se $X_i \sim N(\mu_i, \sigma_i^2)$ são independentes e a_i são constantes reais:

$$X = \sum_{i=1}^{k} a_i X_i \sim N\left(\sum_{i=1}^{k} a_i \mu_i,\ \sum_{i=1}^{k} a_i^2 \sigma_i^2\right) \qquad \text{(Fórmula 16)}$$

3 Pares aleatórios

3.1 Distribuição conjunta, marginais e condicionais

Função de distribuição acumulada (FDA)

Para uma variável aleatória contínua X:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)\, dt \qquad \text{(Fórmula 17)}$$

3.2 Independência

Independência - Caso discreto

X e Y são independentes se, para todos os valores possíveis x e y:

$$P(X = x,\ Y = y) = P(X = x) \cdot P(Y = y) \qquad \text{(Fórmula 18)}$$

Independência - Caso contínuo

X e Y são independentes se, para todos os valores x e y:

$$f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y) \qquad \text{(Fórmula 19)}$$

3.3 Covariância e correlação

Covariância

$$\mathrm{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y) \qquad \text{(Fórmula 20)}$$

Correlação

$$\mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}} \qquad \text{(Fórmula 21)}$$

Interpretação da correlação:

- $\mathrm{Corr}(X, Y) = 0$: Ausência de correlação linear (não implica independência)
- $\mathrm{Corr}(X, Y) > 0$: Associação linear positiva
- $\mathrm{Corr}(X, Y) < 0$: Associação linear negativa
- $|\mathrm{Corr}(X, Y)| = 1$: Correlação linear perfeita
- $|\mathrm{Corr}(X, Y)| \approx 0.24$: Correlação linear fraca

Exemplos de interpretação:

- **Se** $\mathrm{Corr}(X, Y) \neq 0$**:** Uma vez que $\mathrm{Corr}(X, Y) \neq 0$, concluímos que X e Y são variáveis aleatórias dependentes. [É sabido que: caso X e Y sejam v.a. independentes, então $\mathrm{Corr}(X, Y) = 0$.]

- **Se** $\mathrm{Corr}(X, Y) < 0$**:** Dado que $\mathrm{Corr}(X, Y) < 0$, podemos adiantar que X e Y tenderão a variar em sentidos opostos relativamente aos respetivos valores esperados.

- **Se** $|\mathrm{Corr}(X, Y)| \approx 0.24$**:** Como $|\mathrm{Corr}(X, Y)| \approx 0.24$ dista bastante de 1, podemos afirmar que as v.a. estão fracamente correlacionadas. [Importa notar que o coeficiente de correlação quantifica a associação linear entre as v.a. X e Y, logo não captura uma eventual relação não linear entre estas duas v.a.]

4 Combinações lineares e teorema do limite central

4.1 Combinações lineares de variáveis aleatórias

4.2 Distribuição assintótica da soma e da média

Teorema do Limite Central (TLC)

Seja $X_1, X_2, \ldots, X_n$ uma amostra i.i.d. com média μ e variância σ^2. Para n suficientemente grande ($n \geq 30$):

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \overset{a}{\sim} N(0,1) \quad \text{onde } S_n = X_1 + \cdots + X_n \qquad \text{(Fórmula 22)}$$

Equivalentemente, para a média amostral:

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \overset{a}{\sim} N(0,1) \quad \text{onde } \overline{X} = \frac{S_n}{n} \qquad \text{(Fórmula 22a)}$$

Quando usar o TLC:

- **Distribuição original desconhecida ou não-normal:** O TLC permite usar a aproximação normal para a soma ou média amostral, desde que $n \geq 30$.
- **Distribuição original normal:** A soma ou média amostral tem distribuição exatamente normal para qualquer n (não é necessário recorrer ao TLC).
- **Aplicação prática:** Permite calcular probabilidades e construir intervalos de confiança quando a distribuição populacional é desconhecida.

Notas importantes:

- Válido independentemente da distribuição original das X_i
- A convergência é mais rápida se a distribuição original for simétrica
- Para $n < 30$, só se pode usar se a população for aproximadamente normal

5 Estimação pontual

5.1 Estatísticas e estimadores

5.2 Método da máxima verosimilhança

Função de máxima verosimilhança

Para uma amostra $X_1, \ldots, X_n$ com função de densidade $f_X(x)$ dependente do parâmetro μ:

Função de verosimilhança:

$$L(\mu \mid \underline{x}) = \prod_{i=1}^{n} f_X(x_i) \qquad \text{(Fórmula 23)}$$

Função log-verosimilhança:

$$\ell(\mu \mid \underline{x}) = \ln L(\mu \mid \underline{x}) = \sum_{i=1}^{n} \ln f_X(x_i)$$

Estimador de máxima verosimilhança:

$$\hat{\mu} : \begin{cases} \left. \dfrac{\partial \ln L(\mu \mid \underline{x})}{\partial \mu} \right|_{\mu=\hat{\mu}} = 0 & \text{(ponto estacionário)} \\ \left. \dfrac{\partial^2 \ln L(\mu \mid \underline{x})}{\partial \mu^2} \right|_{\mu=\hat{\mu}} < 0 & \text{(ponto de máximo)} \end{cases}$$

6 Estimação intervalar

6.1 Intervalos de confiança para o valor esperado, variância conhecida

IC para média - Variância conhecida

Para população normal ou amostra grande (TLC):

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1) \qquad \text{(Fórmula 24)}$$

6.2 Intervalos de confiança para o valor esperado, variância desconhecida

IC para média - Variância desconhecida (amostra pequena, população normal)

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{(n-1)} \qquad \text{(Fórmula 25)}$$

IC para média - Variância desconhecida (amostra grande)

$$Z = \frac{\overline{X} - \mu}{S/\sqrt{n}} \overset{a}{\sim} N(0, 1) \qquad \text{(Fórmula 26)}$$

6.3 Intervalo de confiança para a variância

IC para variância - População normal, média desconhecida

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)} \qquad \text{(Fórmula 27)}$$

6.4 Intervalo de confiança para probabilidade de sucesso

IC para proporção - População de Bernoulli

Para $X \sim \text{Bernoulli}(p)$ com amostra grande:

$$Z = \frac{\overline{X} - p}{\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}} \sim N(0, 1) \qquad \text{(Fórmula 28)}$$

7 Testes de hipóteses

7.1 Testes para o valor esperado, variância conhecida

Teste para média - Variância conhecida

Para testar $H_0 : \mu = \mu_0$:

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1) \qquad \text{(Fórmula 29)}$$

Rejeita-se H_0 se $|Z| > z_{1-\alpha/2}$ (teste bilateral).

Cálculo do valor-p

Definição: Probabilidade, sob H_0, de obter um valor da estatística tão extremo ou mais extremo que o observado.

- **Testes bilaterais:** valor-p $= 2 \cdot P(Z > |z_{obs}|)$
- **Testes unilaterais:** valor-p $= P(Z > z_{obs})$ ou $P(Z < z_{obs})$
- **Para estatísticas t:** Substituir Z por t e usar distribuição t
- **Decisão:** Compara-se com α para decidir sobre H_0

7.2 Testes para o valor esperado, variância desconhecida

Teste para média - Variância desconhecida

Para testar $H_0 : \mu = \mu_0$ em população normal:

$$t = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} \sim t_{(n-1)} \qquad \text{(Fórmula 30)}$$

Rejeita-se H_0 se $|t| > t_{n-1,1-\alpha/2}$. Para n grande, usar aproximação normal.

7.3 Testes para a variância

Teste para variância - População normal

Para testar hipóteses sobre σ^2 com média desconhecida:

$$T = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)} \qquad \text{(Fórmula 31)}$$

7.4 Testes para probabilidade de sucesso

Teste para proporção - População de Bernoulli

Para testar $H_0 : p = p_0$ com amostra grande ($n > 30$):

$$T = \frac{\overline{X} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \overset{a}{\sim}_{H_0} N(0,1) \qquad \text{(Fórmula 32)}$$

7.5 Teste de ajustamento do qui-quadrado

Teste de ajustamento de Pearson

Para hipótese nula simples:

$$T = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \overset{H_0}{\sim} \chi^2_{(k-1)} \qquad \text{(Fórmula 33)}$$

Onde O_i são frequências observadas, E_i são frequências esperadas e k é o número de categorias.

8 Introdução à regressão linear simples

8.1 Modelo de regressão linear simples

Estimadores dos parâmetros de regressão

Inclinação:

$$\hat{\beta}_1 = \frac{\sum x_i Y_i - n\overline{x}\,\overline{Y}}{\sum x_i^2 - n\overline{x}^2} \qquad \text{(Fórmula 34)}$$

Ordenada na origem:

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1\,\overline{x} \qquad \text{(Fórmula 35)}$$

8.2 Intervalos de confiança e testes para os parâmetros

Estatísticas para inferência sobre β_0

$$\frac{\hat{\beta}_0 - \beta_0}{\sqrt{\left(\frac{1}{n} + \frac{\overline{x}^2}{\sum x_i^2 - n\overline{x}^2}\right)\hat{\sigma}^2}} \sim t_{(n-2)} \qquad \text{(Fórmula 36)}$$

Estatísticas para inferência sobre β_1

$$\frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\hat{\sigma}^2}{\sum x_i^2 - n\overline{x}^2}}} \sim t_{(n-2)} \qquad \text{(Fórmula 37)}$$

Estatísticas para inferência sobre $\beta_0 + \beta_1 x_0$

$$\frac{(\hat{\beta}_0 + \hat{\beta}_1 x_0) - (\beta_0 + \beta_1 x_0)}{\sqrt{\left(\frac{1}{n} + \frac{(\overline{x} - x_0)^2}{\sum x_i^2 - n\overline{x}^2}\right)\hat{\sigma}^2}} \sim t_{(n-2)} \qquad \text{(Fórmula 38)}$$

Teste para valor esperado da resposta em x_0

Para testar $H_0 : E(Y|x_0) = e_0$:

$$T = \frac{\hat{\beta}_0 + \hat{\beta}_1 x_0 - e_0}{\sqrt{\left(\frac{1}{n} + \frac{(\overline{x} - x_0)^2}{\sum_{i=1}^n x_i^2 - n\overline{x}^2}\right)\hat{\sigma}^2}} \sim_{H_0} t_{(n-2)} \qquad \text{(Fórmula 39)}$$

8.3 Coeficiente de determinação

Coeficiente de determinação R^2

Mede a proporção da variabilidade de Y explicada pelo modelo linear:

$$R^2 = \frac{\left(\sum_{i=1}^{n} x_i Y_i - n\overline{x}\,\overline{Y}\right)^2}{\left(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2\right)\left(\sum_{i=1}^{n} Y_i^2 - n\overline{Y}^2\right)} \qquad \text{(Fórmula 40)}$$

Interpretação geral:

- $R^2 \approx 0$: Modelo não explica a variabilidade de Y
- $R^2 \approx 1$: Modelo explica quase toda a variabilidade de Y
- $0 < R^2 < 1$: Proporção da variância explicada pelo modelo

Exemplo de interpretação:

- **Se** $R^2 = 0.34$**:** Cerca de 34% da variação total da variável resposta Y é explicada pela variável x, através do modelo de regressão linear simples ajustado. Podemos concluir que a reta estimada não se ajusta bem ao conjunto de dados.
- **Critério de ajuste:** Valores de R^2 próximos de 0.7 ou superiores indicam um bom ajuste do modelo aos dados.

Formulário completo com 40 fórmulas numeradas (retiradas de exames)

Use-se o índice de fórmulas (página 3) para navegação direta