Distribuições amostrais

Tiago M. Magalhães

Departamento de Estatística - ICE-UFJF

Juiz de Fora, 04 de outubro de 2024

Roteiro

Roteiro

Média amostral

Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias (VA) independentes e identicamente distribuídas (IID), com função distribuição $\mathcal{D}$, valor esperado μ e variância σ^2,

Média amostral

Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias (VA) independentes e identicamente distribuídas (IID), com função distribuição $\mathcal{D}$, valor esperado μ e variância σ^2, isto é, $X_i \sim \mathcal{D}(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$.

Média amostral

Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias (VA) independentes e identicamente distribuídas (IID), com função distribuição $\mathcal{D}$, valor esperado μ e variância σ^2, isto é, $X_i \sim \mathcal{D}(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$. Diz-se que tal sequência de VA constitui uma amostra da distribuição $\mathcal{D}$.

Média amostral

Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias (VA) independentes e identicamente distribuídas (IID), com função distribuição $\mathcal{D}$, valor esperado μ e variância σ^2, isto é, $X_i \sim \mathcal{D}(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$. Diz-se que tal sequência de VA constitui uma amostra da distribuição $\mathcal{D}$.

Média amostral

A quantidade,

$$\bar{X} = \sum_{i=1}^{n} \frac{X_i}{n} \tag{1}$$

é chamada de média amostral.

Média amostral

Notem que,

$$\mathbb{E}\left(\bar{X}\right) = \mathbb{E}\left(\sum_{i=1}^{n} \frac{X_i}{n}\right) = \frac{1}{n}\mathbb{E}\left(\sum_{i=1}^{n} X_i\right)$$
$$= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mu = \mu,$$

Média amostral

Notem que,

$$\mathbb{E}\left(\bar{X}\right) = \mathbb{E}\left(\sum_{i=1}^{n} \frac{X_i}{n}\right) = \frac{1}{n}\mathbb{E}\left(\sum_{i=1}^{n} X_i\right)$$
$$= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left(X_i\right) = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu,$$

isto é, o valor esperado de uma média amostral é μ, a média da distribuição.

Média amostral

Notem que,

$$\begin{aligned}\mathbb{E}\left(\bar{X}\right) &= \mathbb{E}\left(\sum_{i=1}^{n}\frac{X_i}{n}\right) = \frac{1}{n}\mathbb{E}\left(\sum_{i=1}^{n}X_i\right)\\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mu = \mu,\end{aligned}$$

isto é, o valor esperado de uma média amostral é μ, a média da distribuição. Quando a média da distribuição é desconhecida, a média amostral é frequentemente utilizada para estimá-la.

Média amostral

Notem que,

$$\mathbb{E}\left(\bar{X}\right) = \mathbb{E}\left(\sum_{i=1}^{n}\frac{X_i}{n}\right) = \frac{1}{n}\mathbb{E}\left(\sum_{i=1}^{n}X_i\right)$$
$$= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mu = \mu,$$

isto é, o valor esperado de uma média amostral é μ, a média da distribuição. Quando a média da distribuição é desconhecida, a média amostral é frequentemente utilizada para estimá-la.

Média amostral

Adicionalmente,

$$\text{Var}\left(\bar{X}\right) = \text{Var}\left(\sum_{i=1}^{n} \frac{X_i}{n}\right) = \left(\frac{1}{n}\right)^2 \text{Var}\left(\sum_{i=1}^{n} X_i\right)$$

$$\stackrel{\text{Ind.}}{=} \frac{1}{n^2} \sum_{i=1}^{n} \text{Var}\left(X_i\right) = \frac{1}{n^2} \sum_{i=1}^{n} \sigma^2 = \frac{\sigma^2}{n}.$$

Média amostral

Adicionalmente,

$$\text{Var}\left(\bar{X}\right) = \text{Var}\left(\sum_{i=1}^{n} \frac{X_i}{n}\right) = \left(\frac{1}{n}\right)^2 \text{Var}\left(\sum_{i=1}^{n} X_i\right)$$

$$\stackrel{\text{Ind.}}{=} \frac{1}{n^2} \sum_{i=1}^{n} \text{Var}\left(X_i\right) = \frac{1}{n^2} \sum_{i=1}^{n} \sigma^2 = \frac{\sigma^2}{n}.$$

Desvios

As quantidades,

$$X_i - \bar{X},$$

$i = 1, 2, \ldots, n,$

Desvios

As quantidades,

$$X_i - \bar{X},$$

$i = 1, 2, \ldots, n$, são chamadas de desvios e elas são iguais às diferenças entre os valores individuais dos dados e a média amostral.

Desvios

As quantidades,

$$X_i - \bar{X},$$

$i = 1, 2, \ldots, n$, são chamadas de desvios e elas são iguais às diferenças entre os valores individuais dos dados e a média amostral.

Desvios

E nós temos ainda a seguinte relação,

$$
\begin{aligned}
\text{Cov}\left(X_i - \bar{X}, \bar{X}\right) &= \text{Cov}\left(X_i, \bar{X}\right) - \text{Cov}\left(\bar{X}, \bar{X}\right) \\
&= \text{Cov}\left(X_i, \frac{1}{n}\sum_{j=1}^{n} X_j\right) - \text{Var}\left(\bar{X}\right) \\
&= \frac{1}{n}\sum_{j=1}^{n} \text{Cov}\left(X_i, X_j\right) - \frac{\sigma^2}{n} \\
&= \frac{\sigma^2}{n} - \frac{\sigma^2}{n} = 0. \qquad (2)
\end{aligned}
$$

Desvios

E nós temos ainda a seguinte relação,

$$\begin{aligned}
\text{Cov}\left(X_i - \bar{X}, \bar{X}\right) &= \text{Cov}\left(X_i, \bar{X}\right) - \text{Cov}\left(\bar{X}, \bar{X}\right) \\
&= \text{Cov}\left(X_i, \frac{1}{n}\sum_{j=1}^{n} X_j\right) - \text{Var}\left(\bar{X}\right) \\
&= \frac{1}{n}\sum_{j=1}^{n} \text{Cov}\left(X_i, X_j\right) - \frac{\sigma^2}{n} \\
&= \frac{\sigma^2}{n} - \frac{\sigma^2}{n} = 0. \qquad (2)
\end{aligned}$$

Desvios

Embora $\bar{X}$ e o desvio $X_i - \bar{X}$ sejam não correlacionados, eles não são, em geral, independentes.

Desvios

Embora $\bar{X}$ e o desvio $X_i - \bar{X}$ sejam não correlacionados, eles não são, em geral, independentes.

Entretanto, no caso especial em que as VA X_i são VA normais,

Desvios

Embora $\bar{X}$ e o desvio $X_i - \bar{X}$ sejam não correlacionados, eles não são, em geral, independentes.

Entretanto, no caso especial em que as VA X_i são VA normais, $\bar{X}$ é independente de toda a sequência de desvios $X_i - \bar{X}$, $i = 1, 2, \ldots, n$.

Desvios

Embora $\bar{X}$ e o desvio $X_i - \bar{X}$ sejam não correlacionados, eles não são, em geral, independentes.

Entretanto, no caso especial em que as VA X_i são VA normais, $\bar{X}$ é independente de toda a sequência de desvios $X_i - \bar{X}$, $i = 1, 2, \ldots, n$.

Roteiro

Variância amostral

Sejam $X_i \sim \mathcal{D}(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$ e $\bar{X}$, a média amostral, definida em (1). A VA,

$$S^2 = \sum_{i=1}^{n} \frac{\left(X_i - \bar{X}\right)^2}{n-1} \tag{3}$$

é chamada de variância amostral.

Variância amostral

Sejam $X_i \sim \mathcal{D}(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$ e $\bar{X}$, a média amostral, definida em (1). A VA,

$$S^2 = \sum_{i=1}^{n} \frac{\left(X_i - \bar{X}\right)^2}{n-1} \tag{3}$$

é chamada de variância amostral.

Variância amostral

Nós podemos escrever (3) da seguinte maneira:

$$
\begin{aligned}
(n-1)\,S^2 &= \sum_{i=1}^{n}\left(X_i - \mu + \mu - \bar{X}\right)^2 = \sum_{i=1}^{n}\left[(X_i-\mu) - \left(\bar{X}-\mu\right)\right]^2 \\
&= \sum_{i=1}^{n}(X_i-\mu)^2 - 2\left(\bar{X}-\mu\right)\sum_{i=1}^{n}(X_i-\mu) + \sum_{i=1}^{n}\left(\bar{X}-\mu\right)^2 \\
&= \sum_{i=1}^{n}(X_i-\mu)^2 - 2\left(\bar{X}-\mu\right)n\left(\bar{X}-\mu\right) + n\left(\bar{X}-\mu\right)^2 \\
&= \sum_{i=1}^{n}(X_i-\mu)^2 - n\left(\bar{X}-\mu\right)^2. \qquad (4)
\end{aligned}
$$

Variância amostral

Nós podemos escrever (3) da seguinte maneira:

$$
\begin{aligned}
(n-1)\,S^2 &= \sum_{i=1}^{n}\left(X_i-\mu+\mu-\bar{X}\right)^2 = \sum_{i=1}^{n}\left[(X_i-\mu)-\left(\bar{X}-\mu\right)\right]^2 \\
&= \sum_{i=1}^{n}(X_i-\mu)^2 - 2\left(\bar{X}-\mu\right)\sum_{i=1}^{n}(X_i-\mu) + \sum_{i=1}^{n}\left(\bar{X}-\mu\right)^2 \\
&= \sum_{i=1}^{n}(X_i-\mu)^2 - 2\left(\bar{X}-\mu\right)n\left(\bar{X}-\mu\right) + n\left(\bar{X}-\mu\right)^2 \\
&= \sum_{i=1}^{n}(X_i-\mu)^2 - n\left(\bar{X}-\mu\right)^2. \qquad (4)
\end{aligned}
$$

Variância amostral

Calculando a esperança da equação anterior, nós obtemos:

$$
\begin{aligned}
(n-1)\,\mathbb{E}\left(S^2\right) &= \sum_{i=1}^{n} \mathbb{E}\left[(X_i-\mu)^2\right] - n\mathbb{E}\left[\left(\bar{X}-\mu\right)^2\right] \\
&= n\sigma^2 - n\mathrm{Var}\left(\bar{X}\right) = (n-1)\,\sigma^2 \\
\Rightarrow \mathbb{E}\left(S^2\right) &= \sigma^2.
\end{aligned}
$$

Variância amostral

Calculando a esperança da equação anterior, nós obtemos:

$$
\begin{aligned}
(n-1)\,\mathbb{E}\left(S^2\right) &= \sum_{i=1}^{n} \mathbb{E}\left[(X_i-\mu)^2\right] - n\mathbb{E}\left[\left(\bar{X}-\mu\right)^2\right] \\
&= n\sigma^2 - n\mathrm{Var}\left(\bar{X}\right) = (n-1)\,\sigma^2 \\
\Rightarrow \mathbb{E}\left(S^2\right) &= \sigma^2.
\end{aligned}
$$

Logo, o valor esperado da variância amostral é a variância da distribuição $\mathcal{D}$.

Variância amostral

Calculando a esperança da equação anterior, nós obtemos:

$$
\begin{aligned}
(n-1)\,\mathbb{E}\left(S^2\right) &= \sum_{i=1}^{n} \mathbb{E}\left[(X_i-\mu)^2\right] - n\mathbb{E}\left[\left(\bar{X}-\mu\right)^2\right] \\
&= n\sigma^2 - n\mathrm{Var}\left(\bar{X}\right) = (n-1)\,\sigma^2 \\
\Rightarrow \mathbb{E}\left(S^2\right) &= \sigma^2.
\end{aligned}
$$

Logo, o valor esperado da variância amostral é a variância da distribuição $\mathcal{D}$.

Roteiro

Distribuição conjunta da média e da variância amostral

Sejam $X_1, X_2, \ldots, X_n$ VA IID normais com valor esperado μ e variância σ^2.

Seja também $\bar{X}$, a média amostral, definida em (1).

Distribuição conjunta da média e da variância amostral

Sejam $X_1, X_2, \ldots, X_n$ VA IID normais com valor esperado μ e variância σ^2.

Seja também $\bar{X}$, a média amostral, definida em (1). Nós temos que,

$$\bar{X} \sim \text{N}\left(\mu, \frac{\sigma^2}{n}\right). \tag{5}$$

Distribuição conjunta da média e da variância amostral

Sejam $X_1, X_2, \ldots, X_n$ VA IID normais com valor esperado μ e variância σ^2.

Seja também $\bar{X}$, a média amostral, definida em (1). Nós temos que,

$$\bar{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right). \tag{5}$$

Distribuição conjunta da média e da variância amostral

Para provar (5), nós podemos utilizar a função geradora de momentos (FGM). Lembrando que a FGM de $X_i \sim N(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$, é dada por:

$$M_{X_i}(t) = \exp\left\{\mu t + \frac{1}{2}\sigma^2 t^2\right\}.$$

Distribuição conjunta da média e da variância amostral

Para provar (5), nós podemos utilizar a função geradora de momentos (FGM). Lembrando que a FGM de $X_i \sim \mathrm{N}(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$, é dada por:

$$M_{X_i}(t) = \exp\left\{\mu t + \frac{1}{2}\sigma^2 t^2\right\}.$$

Distribuição conjunta da média e da variância amostral

A FGM de $\bar{X}$, $M_{\bar{X}}(t)$, é dada por:

$$\begin{aligned}
M_{\bar{X}}(t) &= \mathbb{E}\left[e^{t\bar{X}}\right] = \mathbb{E}\left[e^{t\sum_{i=1}^{n}\frac{X_i}{n}}\right] = \mathbb{E}\left[e^{\frac{1}{n}tX_1} \times e^{\frac{1}{n}tX_2} \times \cdots \times e^{\frac{1}{n}tX_n}\right] \\
&\overset{\text{Ind.}}{=} \mathbb{E}\left[e^{\frac{1}{n}tX_1}\right] \times \mathbb{E}\left[e^{\frac{1}{n}tX_2}\right] \times \cdots \times \mathbb{E}\left[e^{\frac{1}{n}tX_n}\right] \\
&= M_{X_1}\left(\frac{t}{n}\right) \times M_{X_2}\left(\frac{t}{n}\right) \times \cdots \times M_{X_n}\left(\frac{t}{n}\right) \\
&\overset{\text{I.D.}}{=} \left[M_{X_1}\left(\frac{t}{n}\right)\right]^n = \left\{\exp\left[\mu\left(\frac{t}{n}\right) + \frac{\sigma^2}{2}\left(\frac{t}{n}\right)^2\right]\right\}^n \\
&= \exp\left\{n\left[\mu\left(\frac{t}{n}\right) + \frac{\sigma^2}{2}\left(\frac{t}{n}\right)^2\right]\right\} = \exp\left\{\mu t + \frac{1}{2}\frac{\sigma^2}{n}t^2\right\}.
\end{aligned}$$

Distribuição conjunta da média e da variância amostral

A FGM de $\bar{X}$, $M_{\bar{X}}(t)$, é dada por:

$$\begin{aligned}
M_{\bar{X}}(t) &= \mathbb{E}\left[e^{t\bar{X}}\right] = \mathbb{E}\left[e^{t\sum_{i=1}^{n}\frac{X_i}{n}}\right] = \mathbb{E}\left[e^{\frac{1}{n}tX_1} \times e^{\frac{1}{n}tX_2} \times \cdots \times e^{\frac{1}{n}tX_n}\right] \\
&\overset{\text{Ind.}}{=} \mathbb{E}\left[e^{\frac{1}{n}tX_1}\right] \times \mathbb{E}\left[e^{\frac{1}{n}tX_2}\right] \times \cdots \times \mathbb{E}\left[e^{\frac{1}{n}tX_n}\right] \\
&= M_{X_1}\left(\frac{t}{n}\right) \times M_{X_2}\left(\frac{t}{n}\right) \times \cdots \times M_{X_n}\left(\frac{t}{n}\right) \\
&\overset{\text{I.D.}}{=} \left[M_{X_1}\left(\frac{t}{n}\right)\right]^n = \left\{\exp\left[\mu\left(\frac{t}{n}\right) + \frac{\sigma^2}{2}\left(\frac{t}{n}\right)^2\right]\right\}^n \\
&= \exp\left\{n\left[\mu\left(\frac{t}{n}\right) + \frac{\sigma^2}{2}\left(\frac{t}{n}\right)^2\right]\right\} = \exp\left\{\mu t + \frac{1}{2}\frac{\sigma^2}{n}t^2\right\}.
\end{aligned}$$

Logo, $\bar{X} \sim \mathrm{N}(\mu, \sigma^2/n)$.

A FGM de $\bar{X}$, $M_{\bar{X}}(t)$, é dada por:

$$\begin{aligned}
M_{\bar{X}}(t) &= \mathbb{E}\left[e^{t\bar{X}}\right] = \mathbb{E}\left[e^{t\sum_{i=1}^{n}\frac{X_i}{n}}\right] = \mathbb{E}\left[e^{\frac{1}{n}tX_1} \times e^{\frac{1}{n}tX_2} \times \cdots \times e^{\frac{1}{n}tX_n}\right] \\
&\overset{\text{Ind.}}{=} \mathbb{E}\left[e^{\frac{1}{n}tX_1}\right] \times \mathbb{E}\left[e^{\frac{1}{n}tX_2}\right] \times \cdots \times \mathbb{E}\left[e^{\frac{1}{n}tX_n}\right] \\
&= M_{X_1}\left(\frac{t}{n}\right) \times M_{X_2}\left(\frac{t}{n}\right) \times \cdots \times M_{X_n}\left(\frac{t}{n}\right) \\
&\overset{\text{I.D.}}{=} \left[M_{X_1}\left(\frac{t}{n}\right)\right]^n = \left\{\exp\left[\mu\left(\frac{t}{n}\right) + \frac{\sigma^2}{2}\left(\frac{t}{n}\right)^2\right]\right\}^n \\
&= \exp\left\{n\left[\mu\left(\frac{t}{n}\right) + \frac{\sigma^2}{2}\left(\frac{t}{n}\right)^2\right]\right\} = \exp\left\{\mu t + \frac{1}{2}\frac{\sigma^2}{n}t^2\right\}.
\end{aligned}$$

Logo, $\bar{X} \sim \mathrm{N}(\mu, \sigma^2/n)$.

Distribuição conjunta da média e da variância amostral

Da mesma forma, sejam $X_i \sim \mathrm{N}(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$ e S^2, a variância amostral, como definida em (3). Para encontrar a distribuição de S^2, primeiramente, nós precisamos reescrevê-lo da seguinte maneira, ver (4):

Distribuição conjunta da média e da variância amostral

Da mesma forma, sejam $X_i \sim \mathrm{N}(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$ e S^2, a variância amostral, como definida em (3). Para encontrar a distribuição de S^2, primeiramente, nós precisamos reescrevê-lo da seguinte maneira, ver (4):

$$(n-1)\,S^2 = \sum_{i=1}^{n} (X_i - \mu)^2 - n\left(\bar{X} - \mu\right)^2$$

$$\Rightarrow \frac{1}{\sigma^2}(n-1)\,S^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n} (X_i - \mu)^2 - \frac{1}{\sigma^2} n\left(\bar{X} - \mu\right)^2$$

$$\Rightarrow \frac{(n-1)\,S^2}{\sigma^2} = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2 - \left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^2.$$

Distribuição conjunta da média e da variância amostral

Da mesma forma, sejam $X_i \sim \mathrm{N}(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$ e S^2, a variância amostral, como definida em (3). Para encontrar a distribuição de S^2, primeiramente, nós precisamos reescrevê-lo da seguinte maneira, ver (4):

$$
\begin{aligned}
(n-1)\,S^2 &= \sum_{i=1}^{n} (X_i - \mu)^2 - n\left(\bar{X} - \mu\right)^2 \\
\Rightarrow \frac{1}{\sigma^2}(n-1)\,S^2 &= \frac{1}{\sigma^2}\sum_{i=1}^{n} (X_i - \mu)^2 - \frac{1}{\sigma^2} n\left(\bar{X} - \mu\right)^2 \\
\Rightarrow \frac{(n-1)\,S^2}{\sigma^2} &= \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2 - \left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^2.
\end{aligned}
$$

Distribuição conjunta da média e da variância amostral

Como $X_i \sim \mathrm{N}(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$, $(X_i - \mu)/\sigma \sim \mathrm{N}(0, 1)$, $[(X_i - \mu)/\sigma]^2 \sim$ qui-quadrado, com 1 grau de liberdade (visto em Cálculo de Probabilidades I) e $\sum_{i=1}^{n} [(X_i - \mu)/\sigma]^2 \sim$ qui-quadrado, com n graus de liberdade.

E, como $\bar{X} \sim \mathrm{N}(\mu, \sigma^2/n)$, logo $\sqrt{n}(\bar{X} - \mu)/\sigma \sim \mathrm{N}(0, 1)$ e $[\sqrt{n}(\bar{X} - \mu)/\sigma]^2 \sim$ qui-quadrado, com 1 grau de liberdade.

Distribuição conjunta da média e da variância amostral

Como $X_i \sim \mathrm{N}(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$, $(X_i - \mu)/\sigma \sim \mathrm{N}(0, 1)$, $[(X_i - \mu)/\sigma]^2 \sim$ qui-quadrado, com 1 grau de liberdade (visto em Cálculo de Probabilidades I) e $\sum_{i=1}^{n} [(X_i - \mu)/\sigma]^2 \sim$ qui-quadrado, com n graus de liberdade.

E, como $\bar{X} \sim \mathrm{N}(\mu, \sigma^2/n)$, logo $\sqrt{n}(\bar{X} - \mu)/\sigma \sim \mathrm{N}(0, 1)$ e $[\sqrt{n}(\bar{X} - \mu)/\sigma]^2 \sim$ qui-quadrado, com 1 grau de liberdade.

Distribuição conjunta da média e da variância amostral

Portanto,

$$\frac{(n-1)\,S^2}{\sigma^2} = \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 \sim \text{qui-quadrado}(n-1) \ \left(\chi^2_{n-1}\right). \qquad (6)$$

Distribuição conjunta da média e da variância amostral

Portanto,

$$\frac{(n-1)\,S^2}{\sigma^2} = \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 \sim \text{qui-quadrado}(n-1) \ \left(\chi^2_{n-1} \right). \qquad (6)$$

Observação: Por (2),

Distribuição conjunta da média e da variância amostral

Portanto,

$$\frac{(n-1)\,S^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 \sim \text{qui-quadrado}(n-1)\ \left(\chi^2_{n-1}\right). \qquad (6)$$

Observação: Por (2), $\bar{X}$ e S^2 são VA independentes.

Distribuição conjunta da média e da variância amostral

Portanto,

$$\frac{(n-1)\,S^2}{\sigma^2} = \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 \sim \text{qui-quadrado}(n-1) \ \left(\chi^2_{n-1}\right). \qquad (6)$$

Observação: Por (2), $\bar{X}$ e S^2 são VA independentes.

Distribuição conjunta da média e da variância amostral

Resultado

Sejam $X_1, X_2, \ldots, X_n$ VA IID normais com valor esperado μ e variância σ^2, então a média amostral $\bar{X}$ e a variância amostral S^2 são independentes.

Distribuição conjunta da média e da variância amostral

Resultado

Sejam $X_1, X_2, \ldots, X_n$ VA IID normais com valor esperado μ e variância σ^2, então a média amostral $\bar{X}$ e a variância amostral S^2 são independentes. $\bar{X}$ é uma VA normal de média μ e variância σ^2/n e S^2 é uma VA qui-quadrado com $n-1$ graus de liberdade.

Distribuição conjunta da média e da variância amostral

Resultado

Sejam $X_1, X_2, \ldots, X_n$ VA IID normais com valor esperado μ e variância σ^2, então a média amostral $\bar{X}$ e a variância amostral S^2 são independentes. $\bar{X}$ é uma VA normal de média μ e variância σ^2/n e S^2 é uma VA qui-quadrado com $n-1$ graus de liberdade.

Roteiro

Duas populações

Sejam duas amostras independentes, tais que $X_{11}, X_{12}, \ldots, X_{1n_1}$ são VA IID normais com valor esperado μ_1 e variância σ_1^2 e $X_{21}, X_{22}, \ldots, X_{2n_2}$ são VA IID normais com valor esperado μ_2 e variância σ_2^2. Por (6), nós temos que,

$$Q_1 = \frac{(n_1 - 1)\, S_1^2}{\sigma_1^2} \sim \chi^2_{n_1-1} \text{ e } Q_2 = \frac{(n_2 - 1)\, S_2^2}{\sigma_2^2} \sim \chi^2_{n_2-1},$$

Duas populações

Sejam duas amostras independentes, tais que $X_{11}, X_{12}, \ldots, X_{1n_1}$ são VA IID normais com valor esperado μ_1 e variância σ_1^2 e $X_{21}, X_{22}, \ldots, X_{2n_2}$ são VA IID normais com valor esperado μ_2 e variância σ_2^2. Por (6), nós temos que,

$$Q_1 = \frac{(n_1 - 1)\, S_1^2}{\sigma_1^2} \sim \chi^2_{n_1 - 1} \text{ e } Q_2 = \frac{(n_2 - 1)\, S_2^2}{\sigma_2^2} \sim \chi^2_{n_2 - 1},$$

em que S_1^2 e S_2^2 são as variâncias amostrais das populações 1 e 2, respectivamente.

Duas populações

Sejam duas amostras independentes, tais que $X_{11}, X_{12}, \dots, X_{1n_1}$ são VA IID normais com valor esperado μ_1 e variância σ_1^2 e $X_{21}, X_{22}, \dots, X_{2n_2}$ são VA IID normais com valor esperado μ_2 e variância σ_2^2. Por (6), nós temos que,

$$Q_1 = \frac{(n_1 - 1)\, S_1^2}{\sigma_1^2} \sim \chi^2_{n_1 - 1} \text{ e } Q_2 = \frac{(n_2 - 1)\, S_2^2}{\sigma_2^2} \sim \chi^2_{n_2 - 1},$$

em que S_1^2 e S_2^2 são as variâncias amostrais das populações 1 e 2, respectivamente.

Duas populações

Um resultado muito importante diz que, a razão entre as razões das VA Q_1 e Q_2 por seus respectivos graus de liberdade tem distribuição F de Fisher-Snedecor com $n_1 - 1$ e $n_2 - 1$ graus de liberdade, isto é,

$$\frac{Q_1/(n_1-1)}{Q_2/(n_2-1)} = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim \mathrm{F}(n_1-1, n_2-1). \tag{7}$$

Duas populações

Um resultado muito importante diz que, a razão entre as razões das VA Q_1 e Q_2 por seus respectivos graus de liberdade tem distribuição F de Fisher-Snedecor com $n_1 - 1$ e $n_2 - 1$ graus de liberdade, isto é,

$$\frac{Q_1/(n_1 - 1)}{Q_2/(n_2 - 1)} = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim \mathsf{F}(n_1 - 1, n_2 - 1). \tag{7}$$

A razão (7) e, consequentemente, a distribuição F são frequentemente utilizadas na **análise de variância**.

Duas populações

Um resultado muito importante diz que, a razão entre as razões das VA Q_1 e Q_2 por seus respectivos graus de liberdade tem distribuição F de Fisher-Snedecor com $n_1 - 1$ e $n_2 - 1$ graus de liberdade, isto é,

$$\frac{Q_1/(n_1-1)}{Q_2/(n_2-1)} = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim \mathrm{F}(n_1-1, n_2-1). \tag{7}$$

A razão (7) e, consequentemente, a distribuição F são frequentemente utilizadas na **análise de variância**.

Roteiro

Bibliografia

Meyer, P. L. (2017), *Probabilidade: aplicações à Estatística*, 2 edn, Livros Técnicos e Científicos Editora Ltda, Rio de Janeiro.

Ross, S. (2010), *Probabilidade: um curso moderno com aplicações*, 8 edn, Bookman, Porto Alegre.

Ross, S. (2019), *A first course in probability*, 10 edn. global edn, Pearson Education, Malaysia.

Obrigado!

✉ tiago.magalhaes@ufjf.br

ufjf.br/tiago_magalhaes

Departamento de Estatística, Sala 319