**Variáveis:**

Qualitativos Nominal: a ordem das categorias não tem significado

EX: Género, Cor de olhos, Grupo Sanguíneo…

Qualitativos Ordinal: há uma ordem natural das categorias

Ex: Classe Social (Alta, baixa, Media), Nível da escola (1º, 2º, 3º ciclo) …

Quantitativos Discreta: os valores podem ordenar-se, mas entre dois valores consecutivos não pode existir um valor intermedio (contagens)

Ex: nº cigarros, nº letras

Quantitativas Contínua: pode tomar qualquer valor num certo intervalo (medições)

Ex: Altura, peso

**Tabela Frequências**

Variáveis Qualitativas Nominais - não incluem as frequências acumuladas

Variáveis Qualitativas Ordinais ou Variáveis Quantitativas Discretas (com número pequeno de valores distintos) - incluem as frequências acumuladas

Variáveis Quantitativas Contínuas ou Variáveis Quantitativas Discretas (com número elevado de valores distintos) - Neste caso há a necessidade de agrupar os dados em classes e incluem as frequências acumulada

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| i | xi | ni | Ni | fi | Fi |
| Linha | Valor/Classe | Frequência absoluta (Contagem) | Frequência absoluta acumulada (∑ ni) | Freq relativa  ( | Freq Relativa  Acumulado  ∑( |

**Classes Para Tabela Frequências**

Nº de classes (Regra de Sturges)

Amplitude dos dados

**Var Aleatória discreta**

uma variável aleatória diz-se discreta se pode assumir um número finito ou infinito numerável de valores → associada a contagens

= Função de probabilidade

∑

= Função de distribuição

| |

Símbolos

= valor esperado

= variância =

= Desvio Padrão

**Vars** **Aleatória discreta**

**Uniforme Discreta**

Situações que todos os valores têm a mesma probabilidade de ocorrer

DX = {a, a + 1, a + 2, . . . , b}

Caso o Dx seja inteiro consecutivos como Dx{1,2,3,4,5} (se possível tentar transformar em inteiro consecutivos)

|

Se não

|

**Binomial**

Situações em que há 2 resultados possíveis (sucesso e insucesso)

provas de Bernoulli (número do domínio) probabilidade de sucesso

dbinom

pbinom

qbinom

Aditividade binomial

**Poisson**

Probabilidade de eventos num intervalo de tempo

dpois

ppois

qpois

Aditividade poisson

**Var Aleatória Continua**

Uma variável aleatória diz-se contínua se pode assumir um número infinito não numerável de valores.

é uma e função densidade de probabilidade se satisfaz as seguintes propriedades

pode ser ou

| |

Variância

Desvio Padrão

|

**Vars Aleatórias Continuas**

**Exponencial**

Tempo ou distancia entre ocorrências sucessivas

Falta de memoria =

dexp

pexp )=

qexp

**Uniforme Continua**

Mesmo que a uniforme discreta, mas para valores contínuos

1º elemento domínio último elemento do domínio

|

dunif

punif =

qunif

**NORMAL**

dnorm

pnorm

qnorm

Aditividade da normal

Sejam X1, X2, . . . , Xk variáveis aleatórias independentes com distribuição Normal, isto é

então

Combinação Linear da Normal

Qualquer combinação linear de variáveis aleatórias independentes com distribuição Normal, ainda tem distribuição Normal, isto

**Normal Reduzida**

diz-se que uma variável aleatória contínua Z tem distribuição Normal Reduzida (standard ou padrão) se a variável aleatória Z tem distribuição Normal com os parâmetros = 0 e σ = 1

dnorm

pnorm

qnorm

**Exemplo entre normal e normal reduzida**

Dados acima referidos dados no enunciado

qnorm(0.1) =

**Amplitude Total**

Habitualmente representa-se por A.

Para dados não agrupados, a amplitude total define-se como a diferença entre o maior e o menor valor do conjunto de dados (diferença entre os extremos). Isto e, seja um conjunto de dados com observações,

Para dados agrupados em classes, a amplitude total é a diferença entre o limite superior da última classe e o limite inferior da primeira classe

**Amplitude interquartis**

A amplitude interquartil define-se como a diferença entre o 3 o quartil e o 1 o quartil:

**Coeficiente de Variação**

Para amostra

Para População

**Caracterização da Distribuição de Frequências**

(media = mediana = moda) é simétrica

(moda < mediana < média) é assimétrica positiva

(moda > mediana > média) é assimétrica Negativa

**Medidas de Assimetria**

Simétrica | = → Assimétrica positiva | Assimétrica negativa

**Qui-Quadrado**

dchisq

fchisq

qchisq

**T-Student**

dt

ft

qt

Se X e Y são variáveis aleatórias independentes, e ), ent

**F de Snedecor**

df

ft

qf

Se a v.a. , então

Se as v.a. X e Y são independentes, e , entao, se

Desta forma , então .

**Parametros e estimadores**

|  |  |  |
| --- | --- | --- |
|  | População | Amostras |
| Media |  | *=* mean(amostra) |
| Variância |  | = var(amostra) |
| Desvio padrão |  |  |
| Proporção |  |  |

**Estimaçao Pontual**

|  |  |  |
| --- | --- | --- |
|  | **= Parâmetro** | **= Estimador** |
| Media | = E[x] | = |
| Variancia | V[x] |  |
| Desvio padrao |  |  |
| Proporçao 1 |  |  |
| Binomial |  | = p\* |
| Poisson |  |  |
| Normal |  |  |
| Exponencial |  |  |

1 : q = 1-p , q\* =1 – p\* , p=? assumir p=0.5

**Formular IC**

1º escolhe para (parâmetro a estudar)

2º determinar a distribuição modal (tendo em conta o que sabe sobre a população e amostra)

3º Identificar o IC (deduzir se necessário)

4º determinar o e os quartis

5º Calcular e interpretar o IC

**Sobre IC**

-Menos graus de conf amplitude diminui (com o mesmo nº de elementos)

-Aumentar o nº elementos amplitude diminui (com o mesmo grau de conf)

**Formular testes de hipóteses**

1º Formular as hipóteses e tipo de teste (H0 VS H1)

2º Fixar

3º Definir a D.A e a E.Tobs

4ºTomar Decisão (ver tabela)

5º Fazer Conclusão

**Tipo de testes**

- Distribuições Simétricas (Normal Reduzida(z-test) , T-Student)

- Distribuições Simétricas (Qui-Quadrado e F de Snedecor)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| H0 | H1 | Tipo de teste | Rejeitar H0 RC | Rejeitar H0 P value (Simetricas) | Rejeitar H0 P value (Assimetrica) |
|  | > | Teste Unilateral Direito | ETobs € RC | p-value - P(ET >= ETobs) | |
| < | Teste Unilateral Esq | p-value - P(ET <= ETobs) | |
| != | Teste Bilateral | p-value - 2 x P(ET >= |ETobs|) | 2 × min {P (ET ≤ ETobs), P (ET ≥ ETobs)} |
| >= | < | Teste Unilateral Esq | p-value - P(ET <= ETobs) | |
| <= | > | Teste unilateral direito | p-value - P(ET >= ETobs) | |

**Regias Críticas Hipóteses Paramétricos**

Z.Test

Bilateral

Unilateral Esquerda

Unilateral Direito

T de Student

Bilateral

Unilateral Esquerda

Unilateral Direito

Nota: df = n1+n2-2 se 2 amostras Independentes ou n1-1 1 amostra \*

Qui-Quadrado

Bilateral

Unilateral Esquerda

Unilateral Direito

F de Snedecor

Bilateral [0} ] U , +

Unilateral Esquerda

Unilateral Direito

**Testes de Hipóteses:**

**Paramétricos:**

* Servem para confirmar ou rejeitar um valor hipotético para um θ de uma p.f.
* (Para testes de p pode-se criar as amostras no r utilizado o para o O

**Não Paramétricos:**

**Testes de Afastamento:**

* Servem para testar a hipótese de que uma amostra é de uma população com uma certa distribuição.
* Qui-quadrado ->Precisa de tabelas de contingência
* Kolmogorov-Smirnov
* Lilliefors
* Shapiro-Wilk

**Testes de independência:**

* Servem para verificar se existem ou não independência entre 2 vars.
* Qui-Quadrado
* As tabelas de contingência são bidimensionais

**Testes de Afastamento:**

* Servem para verificar se 2 amostras podem ser consideradas da mesma população (se tem lista iguais)
* Wilconox ->amostras emparelhadas
* Man-whitney-> amostras independentes
* Alternativa para o teste paramétrico
* Md -> mediana de d = y-x