

Lista 1

Aluno: Gustavo Luiz Bispo dos Santos - 117210400

Aluno: Diego Amancio Pereira - 116210716

Data: 07/05/2022

Questão 1

O que é **Distribuição de Frequências**? Explique e apresente um exemplo.

Resposta: é um arranjo de valores que uma ou mais variáveis tomam em uma amostra. Cada entrada na tabela contém a frequência ou a contagem de ocorrências de valores dentro de um grupo ou intervalo específico, e deste modo, a tabela resume a distribuição dos valores da amostra.

Exemplo:

Rank	Grau de confiança	Números
1	Concordo	20
2	Concordo em parte	30
3	não tenho certeza	20
4	Discordo em parte	15
5	Discordo	15

Questão 2

Explique de **forma intuitiva** e ou de **forma formal** o que é uma **Variável Aleatória**. Apresente um (1) ou mais exemplos.

Resposta: Uma variável aleatória é uma variável quantitativa, cujo resultado (valor) depende de fatores aleatórios. Um exemplo de uma variável aleatória é o resultado do lançamento de um dado que pode dar qualquer número entre 1 e 6. Embora possamos conhecer os seus possíveis resultados, o resultado em si depende de fatores de sorte. Uma variável aleatória pode ser uma medição de um parâmetro que pode gerar valores diferentes.

Questão 3

Como podem ser classificadas as Variáveis Aleatórias? Apresente um exemplo de cada tipo

Resposta: As variáveis aleatórias podem ser classificadas em variáveis aleatórias discretas, contínuas e mistas.

Discreta: Uma variável aleatória discreta pode assumir valores que podem ser contados. O lançamento de um dado de seis lados é um exemplo de variável aleatória discreta finita. O dado fornece um valor inteiro em todos os lançamentos, de modo que não existe a possibilidade de ele cair de lado e fornecer um valor fracionário como 2,5555.

Contínua: Uma variável aleatória contínua pode assumir qualquer valor numérico em um determinado intervalo ou série de intervalos. O resultado de lançamento de martelo nas Olimpíadas é um exemplo de

variável aleatória contínua. Sabe-se que os valores do lançamento de martelo atingem a distância máxima de 60 metros e a distância mínima classificatória de 30 metros. Todos os lançamentos poderão assumir uma infinidade de possibilidades dentro no intervalo entre 60 metros e 30 metros, pois sempre existirá uma fração para medir a menor diferença possível entre os lançamentos como 59 metros, 25 centímetros, 12 milímetros e assim por diante.

Mista: Existem situações práticas, em que a variável aleatória pode tanto assumir valores discretos X_1, X_2, X_3, \dots quando assumir todos os valores em um determinado intervalo. Um exemplo de uma variável aleatória mista pode ser um experimento em que uma moeda é lançada e uma roleta é girada se o resultado do lançamento da moeda for cara, X é igual ao valor da roleta. Se o resultado do lançamento da moeda for coroa, X é igual a -1 . Há a probabilidade de essa variável aleatória ter o valor -1 , e de ficar no intervalo $[0, 360]$.

Questão 4

Apresente a definição e um exemplo de **função de probabilidade** (f.p.).

Resposta: é uma função que associa um valor de probabilidade à cada possível ocorrência de uma variável aleatória discreta. Por exemplo, se tomarmos a variável aleatória discreta “resultado de um dado”, as possíveis ocorrências são 1, 2, 3, 4, 5 e 6. Se considerarmos um dado não viciado, a função de probabilidade associará a cada uma destas ocorrências uma probabilidade igual a $\frac{1}{6}$.

Questão 5

O que é uma **Distribuição de Probabilidades**?

Resposta: uma distribuição de probabilidade descreve o comportamento aleatório de um fenômeno dependente do acaso. A distribuição de probabilidade pode modelar incertezas e descrever fenômenos físicos, biológicos, econômicos, entre outros.

Questão 6

Existe alguma semelhança entre **Distribuição de Probabilidades** e **Distribuição de Frequências**? Explique.

Resposta: Não. A distribuição de probabilidades associa uma probabilidade a cada resultado numérico de um experimento, por exemplo, no lançamento de um dado cada face tem a mesma probabilidade de ocorrência de $\frac{1}{6}$. Já a distribuição de frequências é um conjunto de valores que uma ou mais variáveis tomam em uma amostra, ou seja, são dados que servirão de base para cálculos posteriores, enquanto a distribuição de probabilidades já se trata de resultados aplicados.

Questão 7

Descreva situações ou **experimentos aleatórios** em que o fenômeno (dados) a ser estudado/investigado pode ser representado por uma **Distribuição Bernoulli**.

Resposta: Podem ser estudados e aferidos com a Distribuição Bernoulli, qualquer experimentos ou situações que envolvam soluções binárias (0 ou 1) que possam ser resumidos em sucesso ou fracasso, por exemplo: “Acertar uma questão”, “tomar um copo de água”, “assistir uma série”.

Questão 8

Qual ou quais são os parâmetros da **Distribuição Bernoulli** que deve(m) ser conhecidos para que seja possível calcular a probabilidade de algum evento de interesse?

Resposta: Os parâmetros são: o conjunto de dados (exemplo: uma questão objetiva de 5 opções) e o critério probabilístico (exemplo: chutar a questão). Nesse caso, teríamos a probabilidade de sucesso quando acertar e a probabilidade de fracasso quando errar a questão.

Questão 9

Apresente a f.p. da **Distribuição Bernoulli** e dê um exemplo de problema cuja solução requeira o cálculo de probabilidade usando a mesma. Apresente o cálculo manual e usando função(ões) do R.

Resposta: A função de probabilidade f dessa distribuição é

$$f(k; p) = \begin{cases} p, & k = 1 \\ 1 - p, & k = 0 \end{cases}$$

Exemplo: Um aluno responde um questão objetiva de 5 opções. Queremos a probabilidade de, no chute, acertar e errar a questão.

Manualmente:

X : chutar a questão

$$P(S) = P(x = 1) = \frac{1}{5}$$

$$P(F) = P(x = 0) = \frac{4}{5}$$

$$\text{Probabilidade de sucesso } p = \frac{1}{5}$$

$$\text{Probabilidade de fracasso } 1 - p = \frac{4}{5}$$

Em R:

```
x <- seq(0, 1, by = 1) # sucesso ou fracasso
prob <- dbern(x, prob = 0.2) # probabilidade de fracasso de 80% e sucesso de 20%
prob # fracassos e sucesso respectivamente
```

```
## [1] 0.8 0.2
```

Questão 10

Descreva situações ou **experimentos aleatórios** em que o fenômeno (dados) a ser estudado/investigado pode ser representado por uma **Distribuição Binomial**.

Resposta: Utilizamos a Distribuição Binomial quando precisamos achar repetidamente a probabilidade em experimentos de soluções binárias (0 ou 1) resumidas em sucesso ou fracasso, por exemplo: “A probabilidade de acertar 3 questões no chute entre 4 questões totais” ou “a probabilidade de 4, entre 12 peças fabricadas, serem defeituosas”.

Questão 11

Qual ou quais são os parâmetros da **Distribuição Binomial** que deve(m) ser conhecido(s) para que seja possível calcular a probabilidade de algum evento de interesse?

Resposta: Os parâmetros são: o número de sucessos que queremos, o número de repetições e a probabilidade de sucesso.

Questão 12

Apresente a f.p. da **Distribuição Binomial** e dê um exemplo de problema cuja solução requeira o cálculo de probabilidade usando a mesma. Apresente o cálculo manual e usando função(ões) do R.

Resposta:

A função de probabilidade f dessa distribuição é

$$f(k; n, p) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Exemplo: A probabilidade de acertar 3 questões (com 5 alternativas) no chute entre 4 questões totais.

Manualmente:

X : nº acertos totais

k : nº sucessos (3)

n : nº repetições (4)

p : probabilidade de sucesso ($\frac{1}{5}$)

$P(X = 3) = ?$

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$P(X = 3) = \binom{4}{3} 0.2^3 (1 - 0.2)^{4-3}$$

$$P(X = 3) = \frac{4!}{3!(4-3)!} 0.2^3 * 0.8 = 0,0256$$

Probabilidade é de 0.0256

Em R:

```
prob <- dbinom(x = 3, size = 4, prob = 0.2)
prob # probabilidade
```

```
## [1] 0.0256
```

Questão 13

Apresente a definição e um exemplo de **função densidade de probabilidade** (f.d.p.)

Resposta: a função densidade de probabilidade (FDP), ou densidade de uma variável aleatória contínua, é uma função que descreve a verossimilhança de uma variável aleatória tomar um valor dado. A função densidade pode ser obtida a partir da função distribuição acumulada a partir da operação de derivação (quando esta é derivável). A função que define a FDP é uma derivada de Radon–Nikodym:

$$f = \frac{dX_*P}{d\mu}$$

Isto é, f é qualquer função mensurável com a propriedade:

$$Pr[X \in A] = \int_{X^{-1}A} dP = \int_A f d\mu$$

para qualquer conjunto mensurável $A \in \mathcal{A}$.

Exemplo prático: Qual é a probabilidade de que uma bactéria viva exatamente 5 horas? A resposta é de 0%. Muitas bactérias vivem por aproximadamente 5 horas, mas não há nenhuma chance de que qualquer bactéria morra em exatamente 5.000000000 horas.

Em vez disso, poderíamos perguntar: qual é a probabilidade de que a bactéria morra entre 5 horas e 5,01 horas? Vamos dizer que a resposta é de 0,02 (ou seja, 2%). A seguir: qual é a probabilidade de que a bactéria morra entre 5 horas e 5.001 horas? A resposta é provavelmente em torno de 0,002, uma vez que este é um décimo do intervalo anterior. A probabilidade de que a bactéria morra entre 5 horas e 5.0001 horas é provavelmente cerca de 0,0002, e assim por diante.

Questão 14

Descreva situações ou **experimentos aleatórios** em que o fenômeno (dados) a ser estudado/investigado pode ser representado por uma **Distribuição Normal**.

Resposta: Podemos utilizar da Distribuição Normal, principalmente, para modelar fenômenos naturais e para analisar variáveis aleatórias contínuas, como por exemplo “analisar o peso em gramas de 100 recém nascidos”.

Questão 15

Qual ou quais são os parâmetros da **Distribuição Normal** que deve(m) ser conhecido(s) para que seja possível calcular a probabilidade de algum evento de interesse?

Resposta: Os parâmetros são: o valor central μ e a variabilidade σ^2 .

Questão 16

Apresente a f.p. da **Distribuição Normal** e dê um exemplo de problema cuja solução requeira o cálculo de probabilidade usando a mesma. Apresente o cálculo manual e usando função(ões) do R.

Resposta:

A função de probabilidade f dessa distribuição é

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Entretanto, quando temos uma **Normal Padrão**, utilizamos a tabela da normal padrão junto a notação $P(Z \leq z)$

Exemplo: Em um hospital o tempo médio de uma cirurgia X é de 129 minutos com um desvio padrão de 14 minutos. Qual a probabilidade, desse tipo de cirurgia, requerer um tempo maior do que dois desvios-padrão acima da média?

Manualmente:

$$p = 1 - P(Z \leq z)$$

$$p = 1 - P(Z < 2)$$

$$p = 1 - 0.9772$$

$$p = 0.0228$$

Em R:

```
prob <- 1 - pnorm(2)
prob # probabilidade
```

```
## [1] 0.02275013
```

Questão 17

*Diga e descreva resumidamente quais são as principais **Etapas/Fases** de uma **Pesquisa Estatística**. Dê um exemplo.*

Resposta: As fases do método estatístico são:

1. **Definição do problema:** determinar como a recolha de dados pode solucionar um problema;
2. **Planejamento:** elaborar como fazer o levantamento dos dados;
3. **Coleta de dados:** reunir dados após o planejamento do trabalho pretendido, bem como definição da periodicidade da coleta (contínua, periódica, ocasional ou indireta);
4. **Correção dos dados coletados:** conferir dados para afastar algum erro por parte da pessoa que os coletou;
5. **Apuração dos dados:** organização e contagem dos dados;
6. **Apresentação dos dados:** montagem de suportes que demonstrem o resultado da coleta dos dados (gráficos e tabelas);
7. **Análise dos dados:** exame detalhado e interpretação dos dados.