UNIVERSIDADE FEDERAL DE SANTA CATARINA

C E N T R O   T E C N O L Ó G I C O

Departamento de Informática e Estatística

CAMPUS UNIVERSITÁRIO – TRINDADE – CAIXA POSTAL 476, CEP: 88040-900 – FLORIANÓPOLIS – SC – TEL.0XX(48) 3721-949

TAREFA 02 da Unidade II.

Referência: Capítulos (4 a 7) da Unidade II do livro-texto, do livro "O andar do bêbado" e aulas ministradas.

Nome: Jessica Regina dos Santos **Matrícula:** 22100626

Q.1: Com base no Cap. 4 do "Andar do bêbado":

a) **Faça um breve resumo destacando três aspectos que você considera relevante em relação com o que tem sido ministrado na disciplina:**

1º aspecto: Aleatoriedade em eventos extraordinários – O capítulo destaca como a aleatoriedade pode explicar a ocorrência de eventos extraordinários, frequentemente mal interpretados como resultados de causas específicas, quando na verdade podem ser fruto do acaso. Essa ideia é importante em Probabilidade e Estatística, que lida com fenômenos imprevisíveis e distribuições de probabilidade para analisar padrões em eventos aleatórios.

2º aspecto: Viés de confirmação e padrões ilusórios – As pessoas tendem a perceber padrões onde eles não existem, o que leva ao viés de confirmação. Em estatística, isso reforça a necessidade de validação rigorosa de hipóteses e a análise de dados com métodos que controlem esse viés.

3º aspecto: Distribuição de eventos e a lei dos grandes números – O autor explora como eventos se distribuem ao longo do tempo e o impacto da amostragem no aparecimento de padrões, algo diretamente ligado à disciplina, que utiliza a lei dos grandes números para estudar o comportamento de médias em grandes amostras e entender distribuições reais de eventos.

b) Faça um breve comentário sobre a APOSTA DE PASCAL:

A Aposta de Pascal, apresentada no capítulo 4, é uma proposta que utiliza o conceito de expectativa matemática para avaliar a racionalidade de uma vida baseada na fé religiosa. Pascal argumenta que, na incerteza sobre a existência de Deus, escolher viver com piedade oferece um "retorno esperado" imensamente positivo – a promessa de vida eterna –, enquanto o custo associado a essa escolha é finito e limitado. Esse raciocínio exemplifica como princípios de probabilidade e análise de risco podem ser aplicados a decisões complexas, extrapolando o uso da expectativa matemática para justificar escolhas em contextos onde a certeza é inatingível. A abordagem é não apenas relevante para a discussão religiosa, mas também ilustra como essa técnica pode ser aplicada em diversas áreas onde decisões precisam ser tomadas sob incerteza.

Q.2 - Considerando o "problema do aniversário" (no Cap, 4 do "Andar do Bêbado)**:**

1- **A situação pode ser avaliada com base em que modelo probabilístico? Justifique.**

O problema do aniversário pode ser avaliado com base no modelo de probabilidade clássica, em que todas as datas são equiprováveis, e busca-se a probabilidade de um evento específico (neste caso, coincidência de aniversários).

2- **Calcule a probabilidade de que na turma (que tem 50 participantes) dois integrantes façam aniversário no mesmo dia (supondo que todas as datas de aniversário sejam equiprováveis).**

Para calcular essa probabilidade, é mais fácil primeiro calcular a probabilidade de que todos os aniversários sejam em dias diferentes e, em seguida, subtrair esse valor de 1.

I) Total de dias no ano: supondo que não há anos bissextos, há 365 dias.

II) Probabilidade de aniversários diferentes:

- A primeira pessoa pode ter seu aniversário em qualquer um dos 365 dias.
- A segunda pessoa tem 364 opções (para evitar o aniversário da primeira).
- A terceira pessoa tem 363 opções, e assim por diante.

A probabilidade de todos os 50 aniversários serem diferentes é dada por:

$$P(diferentes) = \frac{365}{365} \times \frac{364}{365} \times \frac{363}{365} \times \ldots \times \frac{316}{365}$$

Isso pode ser simplificado usando o fatorial:

$$P(diferentes) = \frac{365!}{(365-50)! \cdot 365^{50}}$$

Logo, a probabilidade de pelo menos dois aniversários serem iguais é:

$$P(pelo\ menos\ um\ igual) = 1 - P(diferentes).$$

Com base no Cap. 5 do “Andar do bêbado”:

Q.3 – Faça um breve resumo destacando três aspectos que você considera relevante em relação com o que tem sido ministrado na disciplina:

1º aspecto: Lei dos grandes números – O capítulo 5 explora como essa lei é essencial para entender o comportamento de médias e probabilidades em grandes amostras, um conceito essencial para as análises de estatística e probabilidade.
2º aspecto: Confiança nos resultados – Discute-se a dificuldade de afirmar probabilidades exatas em sistemas complexos, o que destaca a importância da análise estatística em interpretações.
3º aspecto: Interpretação de probabilidades condicionais – O capítulo 5 também ressalta a interpretação correta de probabilidades condicionais, aspecto fundamental na tomada de decisões e previsão de resultados na prática.

Q.4 - Faça um breve comentário sobre o PROCESSO DE BERNOULLI:

O Processo de Bernoulli caracteriza-se por uma sequência de experimentos independentes, cada um com duas possíveis saídas – "sucesso" ou "falha" – e uma probabilidade constante de sucesso. Esse processo fundamenta a distribuição binomial, amplamente aplicada no cálculo de probabilidades em eventos repetidos e independentes. De modo versátil, o Processo de Bernoulli tem relevância tanto para modelos de jogos de azar quanto para a análise de fenômenos científicos, onde a compreensão de padrões probabilísticos em sequências de eventos discretos é essencial.

Q.5 - Qual é a sua opinião sobre “a falácia do jogador”?

A falácia do jogador revela uma compreensão equivocada sobre a natureza da independência em eventos aleatórios, como lançamentos de moeda ou roletas de cassino. Apesar de parecer intuitivo que uma sequência de “caras” ou “coroas” possa sugerir que o próximo lançamento de moeda será diferente, cada tentativa é, na realidade, independente e possui a mesma probabilidade de ocorrência. Esse equívoco mostra como a intuição humana pode nos enganar em questões probabilísticas, destacando a importância de um entendimento sólido de conceitos estatísticos e probabilísticos. Ao reconhecer e corrigir essas falácias, podemos tomar decisões mais racionais em contextos de incerteza e evitar suposições infundadas em situações de risco.

Q.6 – Uma prova de concurso, de múltipla escolha, contém **"cd"** questões, cada uma com quatro alternativas com apenas uma correta. Você se submete à prova e decide "chutar" todas as questões.

Obs.: cd = refere-se aos 3º e 4º dígitos de sua matrícula!

A situação experimental pode ser avaliada com base no Modelo de Bernoulli? Justifique.

Sim, a situação pode ser avaliada pelo Modelo de Bernoulli. Cada questão tem apenas duas possibilidades: acerto (sucesso) ou erro (fracasso), com uma probabilidade fixa de 0,25 para acerto (pois há apenas uma alternativa correta dentre quatro). As questões são independentes umas das outras, o que caracteriza uma sequência de experimentos de Bernoulli.

Que modelo pode ser aplicado para calcular a probabilidade de acertar ao menos 60% das questões? Justifique e obtenha a probabilidade!

Modelo a aplicar: Distribuição Binomial.
Justificativa: O modelo binomial é adequado pois estamos lidando com um número fixo de tentativas (10 questões), cada uma com uma probabilidade constante de sucesso (0,25) e são independentes entre si. A distribuição binomial calcula a probabilidade de um número específico de sucessos (acertos) em um conjunto de tentativas.
Cálculos: cálculo da probabilidade de acertar ao menos 60% das questões (ou seja, ao menos 6 acertos).
- Temos uma prova de 10 questões (n = 10), e queremos a probabilidade de acertar pelo menos 6 questões (k ≥ 6).
- A probabilidade de acerto por questão (p) é 0,25, e a probabilidade de erro é 0,75 (pois há 1 alternativa correta em 4).
- A probabilidade de obter k acertos em n tentativas usando a distribuição binomial é dada por $P(X = k) \; = \; {}^{n}_{k} p^k (1-p)^{n-k}$

onde ${}^{n}_{k}$ é o coeficiente binomial, calculado como $\frac{n!}{k!(n-k)!}$
- Calcularei P(X ≥ 6), o que significa somar as probabilidades de acertar exatamente 6, 7, 8, 9 e 10 questões. Ou seja, P(X≥6) = P(X=6) + P(X=7) + P(X=8) + P(X=9) + P(X=10). Para cada valor de k de 6 a 10, aplicamos a fórmula da binomial e somamos esses valores para obter a probabilidade total de acertar pelo menos 6 questões.

Vou calcular P(X ≥ 6), que é a soma das probabilidades P(X=6), P(X=7), P(X=8), P(X=9), P(X=10), usando a fórmula da distribuição binomial:

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

onde,

- n = 10 (número total de questões)
- p = 0.25 (probabilidade de acerto em cada questão)
- q = 1−p = 0.75 (probabilidade de erro em cada questão)
- $\binom{n}{k}$ é o coeficiente binomial, $\frac{n!}{k!(n-k)!}$.

Calculando P(X=6):

Para k = 6

$P(X = 6) = \binom{10}{6} (0.25)^6 (0.75)^4 = \frac{10!}{6!(10-6)!} \times (0.25)^6 \times (0.75)^4$

$$P(X = 6) \approx 0.01056$$

Calculando P(X=7):

Para k = 7

$$P(X = 7) = \binom{10}{7} (0.25)^7 (0.75)^3 = \frac{10!}{7!(10-7)!} \times (0.25)^7 \times (0.75)^3$$

$$P(X = 7) \approx 0.00351$$

Calculando P(X=8):

Para k = 8

$$P(X = 8) \ = \ \genfrac{}{}{0pt}{}{10}{8}\,(0.25)^{8}(0.75)^{2} \ = \ \frac{10!}{8!\,(10-8)!} \times (0.25)^{8} \times (0.75)^{2}$$

$$P(X \ = \ 8) \ \approx \ 0.00038$$

Calculando P(X=9):

Para k = 9

$$P(X = 9) \ = \ \genfrac{}{}{0pt}{}{10}{9}\,(0.25)^{9}(0.75)^{1} \ = \ \frac{10!}{9!\,(10-9)!} \times (0.25)^{9} \times (0.75)^{1}$$

$$P(X \ = \ 9) \ \approx \ 0.000028$$

Calculando P(X=10):

Para k = 10

$$P(X = 10) \ = \ \genfrac{}{}{0pt}{}{10}{10}\,(0.25)^{10}(0.75)^{0} \ = \ \frac{10!}{10!\,(10-10)!} \times (0.25)^{10} \times (0.75)^{0}$$

$$P(X \ = \ 9) \ \approx \ 0.00000095$$

Somando todos os valores:

$$P(X \geq 6) = P(X = 6) + P(X = 7) + P(X = 8) + P(X = 9) + P(X = 10)$$

$$P(X \geq 6) \approx 0.01056 + 0.00351 + 0.00038 + 0.000028 + 0.00000095 \approx 0.0145$$

Portanto, a probabilidade de acertar pelo menos 60% das questões (ou seja, 6 ou mais) é aproximadamente 0.0145, ou 1.45%.

Q.7 – De todos os *bits* transmitidos por um canal digital de comunicação **0,X%** são recebidos com erro. Qual é a probabilidade de que no mínimo um (1) bit seja recebida com erro nos próximos mil (1.000) bits transmitidos?

Obs.: **X =** antepenúltimo dígito de sua matrícula! Se o dígito for zero, use o anterior e assim, sucessivamente!

Compare o **Modelo Binomial X Poisson** e obtenha o **erro relativo** para analisar a aproximação.

Cálculos:

Para resolver essa questão, irei:

- Calcular a probabilidade usando o modelo binomial.
- Calcular a probabilidade usando a aproximação pela distribuição de Poisson.
- Calcular o erro relativo entre os dois resultados para avaliar a precisão da aproximação.

No Modelo Binomial, temos:

Número de tentativas (n) = 1000 bits.

Probabilidade de erro em cada bit (p) = 0,006.

Quero calcular a probabilidade de que pelo menos um bit seja recebido com erro entre os 1000 bits transmitidos.

A probabilidade de pelo menos um erro é dada por $P(no\ mínimo\ um\ erro) = 1 - P(nenhum\ erro)$.

Para nenhum erro ocorrer, tenho X = 0, onde X é uma variável binomial $Binomial(n = 1000,\ p = 0{,}006)$.

Assim, $P(X=0) = (1-p)^n = (1-0{,}006)^{1000}$

E, $P(\text{no mínimo um erro}) = 1 - (1-0{,}006)^{1000} = 0.9976$

Quando n é grande e p é pequeno, podemos aproximar a distribuição binomial por uma distribuição de Poisson com parâmetro $\lambda = np$.

Logo, para a questão, $\lambda = 1000 \times 0{,}006 = 6$

Para a distribuição de Poisson, a probabilidade de pelo menos um erro é dada por

$$P(\text{no mínimo um erro}) = 1 - P(X=0)$$

Como $P(X=0) = e^{-\lambda}$

Assim, $P(\text{no mínimo um erro}) = 1 - e^{-6} = 0.9975$

O erro relativo entre o valor exato (binomial) e o valor aproximado (Poisson) é dado por:

$$\text{Erro relativo} = \left|\frac{(P(\text{binomial}) - P(\text{Poisson}))}{P(\text{binomial})}\right| \cdot 100\% = \left|\frac{(0.9976 - 0.9975)}{0.9976}\right| \cdot 100\% = 0.0045\%$$

A aproximação de Poisson é bastante precisa neste caso, com um erro relativo muito pequeno, o que confirma que a distribuição de Poisson é uma boa aproximação para a binomial quando n é grande e p é pequeno.

Q.8 – Existem vários algoritmos computacionais que permitem gerar números *aleatórios* (ou, mais apropriadamente, *pseudoaleatórios*) no intervalo [0; 1], com variável aleatória X com distribuição uniforme (veja no livro-texto o modelo matemático e suas propriedades).

Considere a geração de 10 amostras de 100 observações/números (X_1, X_2, …., X_{100}) e seja $\overline{X}_i$ a i-ésima média aritmética simples das 10 amostras de 100 observações/números. **Use ao menos 2 algarismos significativos para cada observação gerada.**

Sugestão: Use planilha Excel para responder à questão e anexe-a na solução.

OBS	AM1	AM2	AM3	AM4	AM5	AM6	AM7	AM8	AM9	AM10	Média amostral

1											
2											
.....											
100											

Assista esses sites abaixo para esclarecimento:

https://www.youtube.com/watch?v=XAfkh5zo-Wc

https://www.youtube.com/watch?v=LqXnpIn2Uxs

Planilha utilizada em anexo.

Resposta das questões 8.1 até 8.7 no documento em anexo. Script feito em conjunto com a aluna Myllena Correa.

8.1- Qual é o valor esperado de X -> E(X) e a variância de X -> V(X) para a situação?

8.2 – Qual é a probabilidade de X assumir um valor no intervalo [0,47; 0,53]?

8.3 – Construa o **modelo empírico**[1] (com a apuração/contagem para visualização do histograma) para a média amostral (X), usando um critério técnico adequado.

8.4- Calcule a **média das médias** ou valor esperado de E ($\bar{X}$) e a **variância das médias**, V($\bar{X}$)?

8.5 – Qual é a **distribuição de probabilidade da média amostral**, $\bar{X}$?

8.6 – Calcule a probabilidade de que a média amostral, $\bar{X}$, assuma um valor no intervalo [0,47; 0,53]?

8.7- O que V pode concluir?

[1]Construção do Modelo empírico: etapas básicas

1ª etapa: Identificar os **extremos**

Menor $X_{(1)}$ = e Maior $X_{(n)}$ =

2ª etapa: Calcular a amplitude total ou variação total (**"Range"**) dos dados

Range: $R = X_n - X_1 =$

3ª etapa: Calcular o **número de classes (k),** usando um critério técnico adequado

3.1) Raiz de n: $k = \sqrt{n}$ ou 3.2) Fórmula de Sturges: $k = 1 + 3{,}332 * \log n$

4ª etapa: Calcular a **amplitude** de cada classe

$C = R / k$ (Sugestão: arredondar só no valor de C)

5ª etapa: Fazer a **apuração/contagem** (obter o histograma) dos dados.

Obs.: histograma (ver gráfico no livro-texto)

Q.9 – Num concurso público, a pontuação nos exames avalia-se segundo o modelo da Curva de Gauss utilizando o critério abaixo: Os 10% com pontuação de no mínimo 90, consideram-se aprovados para preencher as vagas existentes; Os 20% abaixo do primeiro grupo ficam classificados para a segunda chamada (caso haja desistência no primeiro grupo), tendo pontuação mínima de 75; Os que pontuaram abaixo do limite do segundo grupo consideram-se desclassificados ou reprovados. Calcule a **média** e o **desvio padrão** para a pontuação esperada nos exames do concurso.

Cálculos:

Para resolver esse problema, vamos supor que as pontuações dos candidatos sejam distribuídas normalmente segundo o modelo da Curva de Gauss, ou seja, a distribuição normal N(μ,σ), onde μ é a média e σ o desvio padrão.

Definindo os percentis para os grupos:

- Os candidatos no primeiro grupo têm uma pontuação mínima de 90 e representam os 10% superiores. O percentil associado ao ponto de corte de 90 é o percentil de 90%, que corresponde a uma pontuação de z = 1.28 (o ponto z = 1,28 é exatamente o valor para o qual 90% da área está abaixo dele na curva normal).
- Os candidatos no segundo grupo (20% abaixo do primeiro) têm uma pontuação mínima de 75, o que corresponde a uma faixa até o percentil de 30%. Esse valor de percentil (30%) tem um valor de z = −0.52 (o ponto de corte para o percentil de 30% está 0,52 desvios padrão abaixo da média, refletindo que ele está à esquerda do centro da distribuição).

Equações para resolver o sistema: para calcular μ e σ, utilizamos as pontuações mínimas e os valores de z para as distribuições normais associadas.

Para o primeiro grupo: $90 = \mu + 1.28\sigma$

Para o segundo grupo: $75 = \mu - 0.52\sigma$

Resolvendo o sistema: subtraindo as duas equações, podemos encontrar os valores de μ e σ.

Subtraindo a segunda equação da primeira, para eliminar μ:

$$90 - 75 = (\mu + 1.28\sigma) - (\mu - 0.52\sigma)$$

$$15 = 1.8\sigma$$

$$\sigma = \frac{15}{1.8} = 8.33$$

Substituindo o valor de σ em uma das equações para encontrar μ:

$$90 = \mu + 1.28 \times 8.33$$
$$90 = \mu + 10.67$$
$$\mu = 90 - 10.67 = 79.33$$

Portanto, a média (μ) é aproximadamente 79.33 e o desvio padrão (σ) é aproximadamente 8.33.

Q.10 – Faça um breve resumo dos tópicos relevantes abordados na Unidade II e sua avaliação quanto à relação com a Unidade I.

Na Unidade II, o foco esteve na construção de uma base para a análise de dados usando probabilidade e modelos teóricos, tanto discretos quanto contínuos. Foram abordados conceitos como eventos, variáveis aleatórias, distribuições de probabilidade (binomial, Poisson, normal, entre outras) e suas aplicações, permitindo a estimativa e a modelagem de incertezas presentes nos dados. Além disso, a unidade explorou o uso de distribuições para modelar fenômenos reais, proporcionando um aprofundamento matemático necessário para a análise quantitativa. Essa unidade complementou a Unidade I, onde foram tratados principalmente a descrição e a exploração de dados de forma descritiva e gráfica. Enquanto a Unidade I focou na compreensão inicial dos dados e suas características, a Unidade II avançou para a modelagem e previsão, possibilitando análises mais profundas e suporte para inferências estatísticas.
Em resumo, a Unidade II estabeleceu as ferramentas probabilísticas e matemáticas que permitiram validar e expandir as observações feitas na Unidade I, fortalecendo a análise exploratória com uma fundamentação teórica.