Aluno: Phedro V N Borges
202001747

Questão 1.1

```
# Definir os parâmetros
n <- 10  # tamanho da amostra
p <- 0.1 # probabilidade de sucesso (peça defeituosa)

# a) Uma peça defeituosa
prob_uma_defeituosa <- dbinom(1, size = n, prob = p)

# b) Nenhuma peça defeituosa
prob_nenhuma_defeituosa <- dbinom(0, size = n, prob = p)

# c) Duas peças defeituosas
prob_duas_defeituosas <- dbinom(2, size = n, prob = p)

# d) No mínimo duas peças defeituosas
prob_pelo_menos_duas_defeituosas <- 1 - pbinom(1, size = n, prob = p)

# e) No máximo duas peças defeituosas
prob_no_maximo_duas_defeituosas <- pbinom(2, size = n, prob = p)

# Exibir os resultados
cat("a) Probabilidade de uma peça defeituosa:", prob_uma_defeituosa, "\n")
cat("b) Probabilidade de nenhuma peça defeituosa:", prob_nenhuma_defeituosa, "\n")
cat("c) Probabilidade de duas peças defeituosas:", prob_duas_defeituosas, "\n")
cat("d) Probabilidade de no mínimo duas peças defeituosas:",
prob_pelo_menos_duas_defeituosas, "\n")
cat("e) Probabilidade de no máximo duas peças defeituosas:",
prob_no_maximo_duas_defeituosas, "\n")
```

Explicação:
Eu abordei essa questão usando a linguagem R. Primeiramente, eu defini os parâmetros do problema, como o tamanho da amostra (n = 10) e a probabilidade de uma peça ser defeituosa (p = 0.1). Para calcular a probabilidade de uma peça ser defeituosa (a), eu utilizei a função dbinom, especificando que quero exatamente uma peça defeituosa. No caso de nenhuma peça ser defeituosa (b), eu novamente empreguei a função dbinom, desta vez indicando que quero exatamente zero defeituosas. Para duas peças defeituosas (c), usei dbinom para calcular a probabilidade de exatamente duas peças serem defeituosas. Para no mínimo duas peças defeituosas (d), eu utilizei pbinom para calcular a probabilidade cumulativa de no máximo uma peça não defeituosa e subtrai esse valor de 1. Para no máximo duas peças defeituosas (e), usei pbinom para calcular a probabilidade cumulativa de no máximo duas peças defeituosas. Finalmente, eu exibi os resultados, mostrando as probabilidades calculadas para cada situação. Essa abordagem permitiu analisar diferentes

cenários na linha de produção. Caso tenha mais alguma pergunta ou algo que não tenha ficado claro, estou à disposição para esclarecer!

Questão 1.2

```r
# Dados da nova terapia
dados_nova_terapia <- c(1.5, 1.2, 1.8, 1.3, 1.7, 1.8, 1.9, 2.1, 1.6, 1.8, 1.8, 1.9, 1.5, 1.7, 1.6,
1.9, 2.1, 2.2, 1.7, 1.7)

# Média amostral
media_amostral <- mean(dados_nova_terapia)

# Média sob a hipótese nula
media_nula <- 2

# Desvio padrão da amostra
desvio_padrao_amostral <- sd(dados_nova_terapia)

# Tamanho da amostra
tamanho_amostra <- length(dados_nova_terapia)

# Estatística t
estatistica_t <- (media_amostral - media_nula) / (desvio_padrao_amostral /
sqrt(tamanho_amostra))

# Graus de liberdade
graus_liberdade <- tamanho_amostra - 1

# Valor crítico para um teste bicaudal com alpha = 0.05
valor_critico <- qt(1 - 0.05 / 2, df = graus_liberdade)

# Cálculo do p-valor
p_valor <- 2 * pt(estatistica_t, df = graus_liberdade)

# Exibição dos resultados
cat("Estatística t:", estatistica_t, "\n")
cat("Valor crítico:", valor_critico, "\n")
cat("P-valor:", p_valor, "\n")
```

Explicação:

Formulação das Hipóteses: Eu formulei a hipótese nula (H0) considerando que a média do tempo de recuperação da nova terapia é igual a 2 anos. A hipótese alternativa (H1) considera que a média é diferente de 2 anos. Significância Estabelecida: Defini um nível de significância (α) de 5%. Coleta e Organização dos Dados: Os dados da nova terapia são: 1.5, 1.2, 1.8, 1.3, 1.7, 1.8, 1.9, 2.1, 1.6, 1.8, 1.8, 1.9, 1.5, 1.7, 1.6, 1.9, 2.1, 2.2, 1.7, 1.7. Teste Estatístico Utilizado: Escolhi usar o teste t de Student para médias independentes, já que não conhecemos o desvio padrão populacional. Cálculos Estatísticos: Calculei a estatística t, o valor crítico e o p-valor usando a fórmula apropriada e funções do R. Tomada de Decisão: Ao comparar a estatística t com o valor crítico e o p-valor com o nível de significância, decidi rejeitar ou não a hipótese nula. Conclusão: Com base nos resultados, rejeitei a hipótese nula. Há evidências estatísticas para afirmar que o tempo de recuperação da nova terapia é significativamente diferente de 2 anos em comparação com a terapia tradicional, com um nível de significância de 5%.

Questão 1.3

```
# Dados
dados_cinto <- c(1.1, 1.2, 1.1, 1.3, 1.0, 1.1, 1.2, 1.1, 1.3, 1.4)
dados_sem_cinto <- c(1.0, 1.1, 1.5, 1.2, 1.1, 1.0, 0.5, 1.2, 1.7, 1.2)

# Teste t de duas amostras independentes
resultado_teste <- t.test(dados_cinto, dados_sem_cinto)

# Exibição dos resultados
cat("Estatística t:", resultado_teste$statistic, "\n")
cat("P-valor:", resultado_teste$p.value, "\n")
```

Explicação:

Formulação das Hipóteses: Eu iniciei formulando as hipóteses estatísticas. A hipótese nula (H0) afirmava que a média do tempo de internação para crianças que usam cinto seria igual à média para aquelas que não usam cinto. A hipótese alternativa (H1) considerava que essas médias seriam diferentes. Significância Estabelecida: Defini um nível de significância de 5% ($\alpha = 0.05$), que é comumente usado em testes de hipóteses. Coleta e Organização dos Dados: Reuni os dados das crianças que usam cinto e aquelas que não usam cinto. Cada conjunto tinha 10 observações. Teste Estatístico Utilizado: Escolhi utilizar o teste t de Student para duas amostras independentes, pois queria comparar as médias de dois grupos distintos. Cálculos Estatísticos: Calculei a estatística t e o p-valor usando a linguagem R. Esses valores são essenciais para a tomada de decisão no teste de hipóteses. Tomada de Decisão: Comparei o p-valor com o nível de significância estabelecido. Se o p-valor fosse menor que 0.05, decidiria rejeitar a hipótese nula. Conclusão: Com base nos resultados obtidos, decidiria se há evidências estatísticas para concluir se o uso do cinto apresenta um tempo médio de internação diferente daqueles que não usam cinto.

Questão 1.4
```
# Dados
tempo_anterior <- c(43, 43, 43, 43, 43, 43, 43, 43, 43, 43,
                    43, 43, 43, 43, 43, 43, 43, 43, 43, 43)

tempo_nova_formula <- c(40.1, 42.2, 38.1, 39.3, 44.0, 42.1, 39.2, 40.1, 40.3, 38.4,
                        39.0, 41.1, 38.5, 39.2, 41.1, 38.0, 40.5, 41.2, 42.7, 45.2)

# Teste t de Student para uma amostra
resultado_teste <- t.test(tempo_nova_formula, mu = mean(tempo_anterior))

# Exibição dos resultados
resultado_teste
```

Explicação:
Eu coletei dados de um novo analgésico e comparei com os dados antigos usando um teste estatístico no R. Os tempos médios de efeito para as duas fórmulas foram analisados para determinar se a nova fórmula é melhor, pior ou igual à anterior. O teste de hipóteses foi realizado com um nível de significância de 5%. Os resultados do teste foram obtidos e podem indicar se há diferença significativa entre as fórmulas.

Questão 1.5
```
# Dados
colesterol_inicio <- c(201, 231, 221, 260, 228, 237, 326, 235, 240, 267, 284, 201)
colesterol_final <- c(200, 236, 216, 233, 224, 216, 296, 195, 207, 247, 210, 209)

# Teste t pareado (amostras relacionadas)
resultado_teste <- t.test(colesterol_inicio, colesterol_final, paired = TRUE)

# Exibição dos resultados
resultado_teste
```

Explicação:
Eu utilizei a linguagem de programação R para analisar os níveis de colesterol de 12 participantes antes e depois de um programa de dieta combinado com exercícios físicos. Realizei um teste estatístico para verificar se houve uma redução significativa nos níveis de colesterol após a implementação do programa. O teste foi feito com um nível de significância de 5%. Os resultados do teste forneceram informações sobre a efetividade do programa na redução do colesterol.

Questão 1.6
```
# Dados
pesos_antes <- c(635, 704, 662, 560, 603, 745, 698, 575, 633, 669)
pesos_depois <- c(640, 712, 681, 558, 610, 740, 707, 585, 635, 682)

# Teste t pareado (amostras relacionadas)
resultado_teste <- t.test(pesos_antes, pesos_depois, paired = TRUE)
```

Exibição dos resultados
resultado_teste

Explicação:
Utilizei o R para realizar uma análise estatística dos pesos de cobaias antes e depois de um tratamento com uma determinada ração. A ideia era verificar se o uso dessa ração contribuiu para o aumento do peso médio dos animais. Fiz um teste estatístico chamado teste t pareado, que compara as duas situações para cada cobaia. A hipótese nula considera que não há diferença, enquanto a hipótese alternativa sugere um aumento significativo. O resultado do teste forneceu um valor-p, e, se esse valor for menor que 0.01 (1% de significância), concluímos que há evidências estatísticas para afirmar que a ração contribuiu para o aumento do peso médio dos animais.

Questão 1.7
```
# Dados
placebo <- c(6, 33, 21, 26, 10, 29, 33, 29, 37, 15, 2, 21, 7, 26, 13)
tianeptina <- c(10, 8, 17, 4, 17, 14, 9, 4, 21, 3, 7, 10, 29, 13, 14, 2)

# Teste t independente
resultado_teste <- t.test(placebo, tianeptina)

# Exibição dos resultados
resultado_teste
```

Explicação:
Utilizei a linguagem de programação R para realizar um teste estatístico comparando os escores de depressão de dois grupos de pacientes: um que recebeu placebo e outro que utilizou Tianeptina. O objetivo era verificar se a média de depressão entre esses grupos é diferente. O teste utilizado foi o teste t independente, e o resultado, incluindo o valor-p, foi apresentado. Se o valor-p for menor que 0.05 (5% de significância), podemos concluir que há evidências estatísticas de diferença nas médias de depressão entre os grupos.

Questão 1.8
```
# Dados
marca_A <- c(17, 20, 23, 20, 18, 19, 20, 18, 19)
marca_B <- c(18, 20, 21, 22, 24, 23, 24, 25, 22, 26, 24)

# Teste t independente
resultado_teste <- t.test(marca_A, marca_B)

# Exibição dos resultados
resultado_teste
```

Explicação:
Utilizei a linguagem de programação R para realizar um teste estatístico comparando os níveis médios de nicotina em duas marcas de cigarros, denominadas A e B. Os resultados dos testes foram apresentados em um teste t independente, que analisa se há diferença significativa nos níveis médios de nicotina entre as marcas. O resultado, incluindo o valor-p, foi exibido, e se o valor-p for menor que 0.05 (5% de significância), podemos concluir que existem evidências estatísticas de diferença nos níveis médios de nicotina entre as duas marcas de cigarros.

Questão 1.9

```
# Dados
antes_cura <- c(224, 270, 400, 444, 590, 660, 1400, 680)
apos_cura <- c(116, 96, 239, 329, 437, 597, 689, 576)

# Teste t pareado
resultado_teste <- t.test(antes_cura, apos_cura, paired = TRUE)

# Exibição dos resultados
resultado_teste
```

Explicação:
Utilizei a linguagem de programação R para conduzir um teste estatístico, especificamente um teste t pareado, para avaliar se há diferenças significativas nos níveis de resíduos de ácido sórbico em peças de presunto antes e após o processo de cura. Os resíduos antes e após o período de cura foram fornecidos como dados. O resultado do teste, incluindo o valor-p, foi apresentado, e ao comparar o valor-p com um nível crítico de 5%, podemos determinar se existem evidências estatísticas de diferença nos níveis de resíduos de ácido sórbico após o processo de cura.

Questão 1.10

```
# Dados
C_I <- c(46307, 55013, 44683, 54528, 48492, 42230, 46709, 50983, 46792, 65775, 57932,
47152)
Cliente <- c(24709, 28023, 21511, 23487, 29644, 21204, 24089, 25958, 26182, 37272,
33650, 25482)
OEM <- c(32007, 33675, 35761, 33987, 31626, 32564, 33078, 34021, 34123, 32347, 33690,
32896)

# Teste ANOVA
resultado_teste <- anova(lm(cbind(C_I, Cliente, OEM) ~ 1))

# Exibição dos resultados
resultado_teste
```

Explicação:

Eu usei a linguagem de programação R para realizar um teste estatístico chamado ANOVA (Análise de Variância) com o objetivo de verificar se as médias de pedidos são iguais para diferentes canais de distribuição: C&I, Cliente e OEM. Os dados foram inseridos no código, e o teste ANOVA foi aplicado aos três conjuntos de dados correspondentes a cada canal. O resultado do teste, incluindo os valores-p, foi apresentado, e a interpretação dos resultados pode ser feita comparando os valores-p com um nível de significância predefinido (geralmente 0,05). Se o valor-p for menor que esse limite, concluímos se há diferenças estatisticamente significativas nas médias de pedidos entre os canais.

Questão 1.11

```
# Dados
varejista1 <- c(218, 188, 225, 217, 176, 187, 221, 212, 210, 203, 188, 185)
varejista2 <- c(101, 87, 123, 101, 95, 97, 93, 131, 76, 101, 87, 114)
varejista3 <- c(268, 296, 321, 312, 301, 294, 285, 305, 289, 303, 324, 332)

# Teste de ANOVA
resultado_anova <- aov(c(varejista1, varejista2, varejista3) ~ rep(1:3, each = 12))

# Resumo do teste ANOVA
summary(resultado_anova)
```

Explicação:
Eu escrevi um código em R para realizar um teste de ANOVA com o objetivo de comparar as médias de demanda mensal entre três varejistas diferentes. No código, organizei os dados de cada varejista em vetores separados e, em seguida, apliquei o teste de ANOVA aos dados combinados dos três varejistas. O resultado do teste inclui estatísticas importantes, como a estatística F e o valor p. A interpretação desses valores pode indicar se há diferenças estatisticamente significativas nas médias de demanda entre os varejistas. Esse tipo de teste é útil para avaliar se os varejistas têm um desempenho significativamente diferente em termos de demanda mensal.

Questão 1.12

```
# Dados observados
tempos <- c(6.8, 7.1, 5.9, 7.5, 6.3, 6.9, 7.2, 7.6, 6.6, 6.3)

# Média e desvio padrão dos tempos observados
media_obs <- mean(tempos)
desvio_padrao_obs <- sd(tempos)

# Tempo médio esperado sob a hipótese nula (sem mudança)
media_esperada <- 7.4

# Número de observações
n <- length(tempos)
```

```
# Estatística de teste (t)
t_stat <- (media_obs - media_esperada) / (desvio_padrao_obs / sqrt(n))

# Graus de liberdade
graus_liberdade <- n - 1

# Valor p (bilateral)
p_valor <- 2 * pt(-abs(t_stat), df = graus_liberdade)

# Teste de hipótese
if (p_valor < 0.05) {
  mensagem <- "Rejeitamos a hipótese nula. Há evidências de que o tempo médio de transmissão foi alterado."
} else {
  mensagem <- "Não temos evidências suficientes para rejeitar a hipótese nula. Não há indícios de alteração no tempo médio de transmissão."
}

# Imprimir resultado
cat("Estatística t:", t_stat, "\n")
cat("Valor p:", p_valor, "\n")
cat(mensagem, "\n")
```

Explicação:

No código em R que forneci, realizei um teste estatístico para verificar se houve uma mudança significativa no tempo médio de transmissão de dados na rede de computadores. Primeiramente, utilizei os tempos de transmissão observados em 10 ensaios independentes. Calculei a média e o desvio padrão desses tempos.

A hipótese nula foi estabelecida considerando que o tempo médio não havia mudado, sendo igual a 7,4 segundos, conforme a média original. A estatística de teste (t) foi calculada para comparar o tempo médio observado com o esperado sob a hipótese nula.

Em seguida, determinei o valor p, que é a probabilidade de observar uma estatística de teste tão extrema quanto a que foi calculada, assumindo que a hipótese nula seja verdadeira. Se o valor p for menor que 0,05, rejeitamos a hipótese nula, indicando evidências de uma alteração no tempo médio de transmissão. Caso contrário, não há indícios suficientes para rejeitar a hipótese nula, sugerindo que o tempo médio permanece inalterado.

Esse teste ajuda a avaliar se as mudanças implementadas na rede tiveram um impacto estatisticamente significativo no tempo de transmissão de dados.

Questão 1.13

```
# Dados
catalisador_A <- c(45, 51, 50, 62, 43, 42, 53, 50, 48, 55)
catalisador_B <- c(45, 35, 43, 59, 48, 45, 41, 43, 49, 39)

# Teste t para duas amostras independentes
resultados_teste <- t.test(catalisador_A, catalisador_B)

# Exibir os resultados
resultados_teste
```

Explicação:

No código em R que forneci, realizei um teste estatístico para verificar se há diferença significativa nos rendimentos entre dois catalisadores, A e B, usando o teste t para duas amostras independentes. Os dados dos rendimentos para cada catalisador foram inseridos como vetores (`catalisador_A` e `catalisador_B`). O resultado do teste, que inclui a estatística t, o valor p e o intervalo de confiança, foi armazenado na variável `resultados_teste`.

Ao interpretar os resultados, focamos especialmente no valor p. Se o valor p for menor que o nível de significância (5% neste caso), isso nos dá evidências estatísticas para rejeitar a hipótese nula, indicando que há diferença significativa nas médias dos rendimentos dos catalisadores A e B. Se o valor p for maior, não teríamos evidências suficientes para rejeitar a hipótese nula, sugerindo que as médias podem não ser significativamente diferentes.

Questão 1.14

```
# Dados
marca_A <- c(33, 38, 36, 40, 31, 35)
marca_B <- c(32, 40, 42, 38, 30, 34)
marca_C <- c(31, 37, 35, 33, 34, 30)
marca_D <- c(28, 34, 32, 30, 33, 31)

# Teste ANOVA
resultado_anova <- aov(c(marca_A, marca_B, marca_C, marca_D) ~ rep(c("A", "B", "C",
"D"), each = 6))

# Sumário do teste
summary(resultado_anova)
```

Explicação:

Neste código em R, eu conduzi um teste estatístico chamado Análise de Variância (ANOVA) para avaliar se existem diferenças significativas entre as médias de quatro tipos de pneus (A, B, C e D). Organizei os dados em vetores correspondentes a cada tipo de pneu e, em seguida, apliquei o teste ANOVA aos dados.

O resultado do teste ANOVA fornece estatísticas importantes, incluindo a estatística F e o valor de p. O valor de p é fundamental para interpretação: se for menor que o nível de significância escolhido (5% neste caso), teríamos evidências estatísticas para rejeitar a hipótese nula. Isso sugeriria que há diferenças significativas nas médias dos pneus. Por outro lado, se o valor de p for maior que 5%, não teríamos evidências suficientes para rejeitar a hipótese nula, indicando que as médias podem não ser significativamente diferentes entre os grupos.

Questão 1.15

```
# Dados
tempo_para_deixar_vicio <- c(15.7, 13.2, 22.6, 13, 10.7, 18.1, 14.7, 7.0, 17.3, 7.5,
                             21.8, 12.3, 19.8, 13.8, 16, 15.5, 13.1, 20.7, 15.5, 9.8)

# Teste t para uma amostra
resultado_teste <- t.test(tempo_para_deixar_vicio, mu = 15)

# Exibir o resultado
resultado_teste
```

Explicação:
No código em R que eu forneço, estou conduzindo um teste estatístico para avaliar se o tempo médio que os indivíduos levam para deixar totalmente o vício do fumo é significativamente diferente de 15 anos. Utilizei o teste t para uma amostra, uma técnica estatística comum para comparar médias.

O vetor "tempo_para_deixar_vicio" contém os dados sobre o tempo que cada indivíduo levou para deixar o vício. O teste é conduzido assumindo que a média populacional é 15 anos (mu = 15).

O resultado do teste é armazenado na variável "resultado_teste". Este resultado inclui estatísticas como a estatística t calculada e o valor de p associado. O valor de p é crucial; se for menor que 0,05 (um nível de significância comum), podemos concluir que existem evidências estatísticas para acreditar que o tempo médio é diferente de 15 anos. Caso contrário, não teríamos evidências suficientes para rejeitar a ideia de que o tempo médio é realmente 15 anos.