1. **Problem**
Uma máquina industrial enche pacotes de café. O peso de cada pacote individual varia seguindo uma distribuição desconhecida, mas um longo histórico de produção indica que a média de peso (μ) é de **483g** e o desvio padrão (σ) é de **11g**.

Se um inspetor de qualidade selecionar uma amostra aleatória de **64** pacotes, qual é a probabilidade percentual de que o **peso médio** dessa amostra seja **maior que 487.5g**?

(Forneça a resposta em porcentagem, com 2 casas decimais. Ex: 5.29)

Solution
O Teorema Central do Limite (TCL) nos permite aproximar a distribuição da média amostral ($\bar{X}$) por uma distribuição Normal.

(a) **Definir a Distribuição da Média Amostral:** Pelo TCL, $\bar{X}$ segue aproximadamente uma distribuição Normal com:
- Média: $E(\bar{X}) = \mu = 483$
- Erro Padrão (desvio padrão da média): $SE(\bar{X}) = \frac{\sigma}{\sqrt{n}} = \frac{11}{\sqrt{64}} \approx 1.375$

(b) **Padronizar o Valor de Interesse (Calcular o Z-score):** Convertemos o valor de 487.5g para a escala da Normal Padrão:

$$Z = \frac{\bar{X} - \mu}{SE(\bar{X})} = \frac{487.5 - 483}{1.375} \approx 3.2727$$

(c) **Calcular a Probabilidade:** A pergunta $P(\bar{X} > 487.5)$ se torna $P(Z > 3.27)$. Usando a função `pnorm()` do R, calculamos a área à direita de Z: `pnorm(3.2727, lower.tail = FALSE)` resulta em $5 \times 10\text{-}4$.

Convertendo para porcentagem, a probabilidade é de **0.05%**.

2. **Problem**
O Teorema Central do Limite é um dos resultados mais importantes da estatística, mas sua aplicação depende de certas condições. Qual das seguintes alternativas **NÃO** é uma condição necessária para a versão clássica do TCL?

(a) A amostra deve ser grande o suficiente.
(b) A população original deve ter média e variância finitas.
(c) As observações na amostra devem ser independentes e identicamente distribuídas (i.i.d.).
(d) A população original deve seguir uma distribuição Normal.

Solution
A resposta correta é que a população original **não precisa** seguir uma distribuição Normal. O poder e a beleza do Teorema Central do Limite residem justamente no fato de que ele é válido para populações com qualquer distribuição (desde que tenham média e variância finitas). As outras três opções são condições necessárias para o teorema.

3. **Problem**
Uma livraria especializada em itens raros recebe, em média, **20 clientes por dia**. O número de clientes que chegam a cada dia é independente dos outros dias.

Considerando um período de **250 dias** de funcionamento, qual é a probabilidade aproximada de que o **número total de clientes** recebidos nesse período seja **menor que 5137**?

(Use a aproximação Normal via Teorema Central do Limite e forneça a resposta em formato decimal com 4 casas. Ex: 0.8765)

Solution
Este é um problema de duas camadas. Primeiro, reconhecemos que o número de clientes

por dia, sendo um evento de contagem em um intervalo fixo, pode ser modelado por uma distribuição de Poisson com $\lambda = 20$. Em uma Poisson, a média e a variância são iguais a λ.

Em seguida, queremos encontrar a distribuição do **número total de clientes** em 250 dias. Seja T o total de clientes, temos $T = X_1 + X_2 + \ldots + X_{250}$, onde cada X_i é uma variável aleatória Poisson($\lambda = 20$) independente e identicamente distribuída.

Pelo **Teorema Central do Limite**, a soma de um grande número de variáveis aleatórias i.i.d. pode ser aproximada por uma distribuição Normal.

(a) **Parâmetros da Distribuição Normal:**
   - **Média da Soma:** $E(T) = n \times \lambda = 250 \times 20 = 5000$
   - **Variância da Soma:** $\text{Var}(T) = n \times \lambda = 250 \times 20 = 5000$
   - **Desvio Padrão da Soma:** $\sigma_T = \sqrt{5000} \approx 70.7107$ Portanto, $T \approx N(5000, 5000)$.

(b) **Calcular a Probabilidade:** Queremos encontrar $P(T < 5137)$. Como estamos usando uma distribuição contínua (Normal) para aproximar uma discreta (Poisson), aplicamos uma **correção de continuidade**, usando $P(T \leq 5136)$, que na aproximação se torna $P(T < 5136.5)$.

   Padronizamos o valor corrigido:

$$Z = \frac{(5137 - 0.5) - 5000}{70.7107} \approx 1.9304$$

   A probabilidade é $P(Z < 1.9304)$, que pode ser encontrada com `pnorm(1.9304)`. O resultado é aproximadamente **0.9732**.

4. **Problem**
   Suponha que a média amostral de 100 observações seja $\bar{X}_n = 15$ e a variância amostral seja $s^2 = 27$. Usando o Método Delta, estime a variância aproximada da transformação logarítmica da média, $\text{Var}(\log(\bar{X}_n))$.

   Solution
   A função de transformação é $g(\mu) = \log(\mu)$. A derivada é $g'(\mu) = 1/\mu$. A fórmula do Método Delta para a variância é: $\text{Var}(g(\bar{X}_n)) \approx [g'(\mu)]^2 \cdot \text{Var}(\bar{X}_n)$.

   Estimamos os parâmetros com os valores da amostra: 1. **Estimativa da derivada:** $g'(\bar{x}) = 1/15 \approx 0.0667$. 2. **Estimativa da variância da média:** $\text{Var}(\bar{X}_n) = s^2/n = 27/100 \approx 0.27$. 3. **Variância via Método Delta:** $(0.0667)^2 \times 0.27 = 0.0012$.

5. **Problem**
   Em um estudo, o Grupo A teve 66 sucessos em 100 tentativas, enquanto o Grupo B teve 49 sucessos em 100 tentativas. O Odds Ratio (OR) é usado para comparar as chances de sucesso entre os grupos. Usando o Método Delta para encontrar o erro padrão do logaritmo do OR, qual é o intervalo de confiança de 95% aproximado para o Odds Ratio?

   (a) [0.8, 2.5]

   (b) [0.95, 1.05]

   (c) [1.64, 4.07]

   (d) [1.14, 3.57]

   Solution

   (a) **Proporções:** $\hat{p}_A = 66/100 = 0.66$ e $\hat{p}_B = 49/100 = 0.49$.
   (b) **Odds Ratio (OR):** $OR = (\hat{p}_A/(1 - \hat{p}_A))/(\hat{p}_B/(1 - \hat{p}_B)) \approx 2.02$.

(c) **Variância do log(OR) (via Método Delta):**

$$\text{Var}(\log(\widehat{OR})) \approx \frac{1}{n_A\hat{p}_A(1-\hat{p}_A)} + \frac{1}{n_B\hat{p}_B(1-\hat{p}_B)} \approx 0.0846$$

(d) **Intervalo de Confiança:** Construímos o IC de 95% para o log(OR) e depois aplicamos a exponencial para obter o IC do OR. O resultado é [1.14, 3.57]. Como este intervalo não contém o valor 1, o resultado é estatisticamente significativo.