

Tópicos especiais em Estatística Aplicada - 2023/2

EPC3

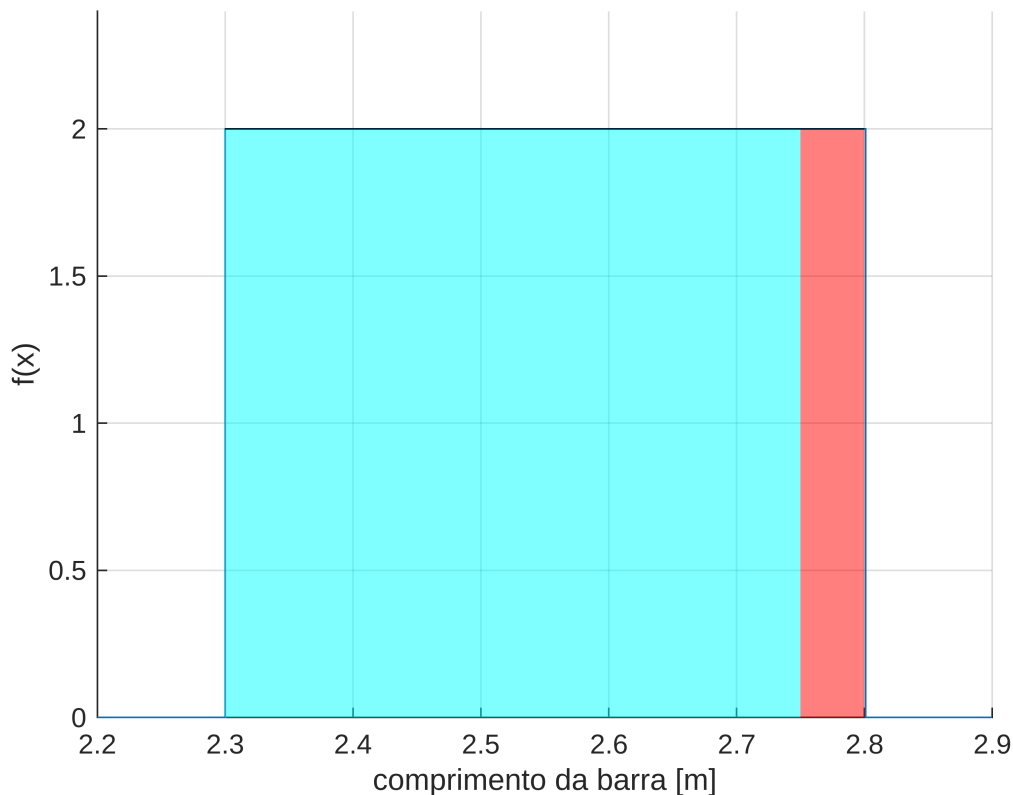
Aluno: João Gabriel Santos Custodio

1) A função densidade de probabilidade do comprimento de uma barra de metal é dada por $f(x) = 2$ para $2.3 < x < 2.8$ m.

1.1. Se as especificações para produzir esta barra forem comprimento de 2.25 a 2.75 m, qual a proporção de barras que não atenderá as especificações?

Para o intervalo de $2.75 < x < 2.8$ as barras não atenderão as especificações. A proporção de barras que não atenderão as especificações pode ser encontrada sendo a área sob a curva desse intervalo de probabilidade.

```
x = linspace(2.2, 2.9, 700);
prob = 2;
figure;
grid(), hold on;
y1 = cat(2, cat(2, zeros(1,100), ones(1,500)*prob), zeros(1,100));
stairs(x, y1)
ylim([0, 1.2*prob])
area([2.3, 2.75], [2, 2], 'FaceColor', 'cyan', 'FaceAlpha', 0.5);
area([2.75, 2.8], [2, 2], 'FaceColor', 'red', 'FaceAlpha', 0.5);
ylabel("f(x)")
xlabel("comprimento da barra [m]")
```



Portanto, evidencia-se que a área sob o gráfico onde não atende as especificações é a área destacada em vermelho no gráfico acima. Dessa forma, temos:

$$(2.8 - 2.75) \cdot 2 = 0.05 \cdot 2 = 0.10 = 10\%.$$

Logo 10% das barras produzidas não atenderão as especificações.

1.2. Suponha que a função densidade de probabilidade seja $f(x) = 2$ para o intervalo de 0.5 m de comprimento. Sobre qual valor deveria $f(x)$ ser centrada para obter a máxima proporção de barras dentro da especificação?

Utilizando o intervalo de 0.5m de comprimento, basta centralizar $f(x)$ no centro do intervalo especificado no item 1.1 e encontrar a mediana. Nesse caso, como já mencionado, o intervalo é de $2.3 < x < 2.75$.

```
x1_2 = linspace(2.3, 2.75, 1000);
median(x1_2)
```

```
ans = 2.5250
```

Conclui-se, portanto, que para nenhuma barra ser produzida fora do intervalo especificado, $f(x)$ há de ser centrada no valor 2.53.

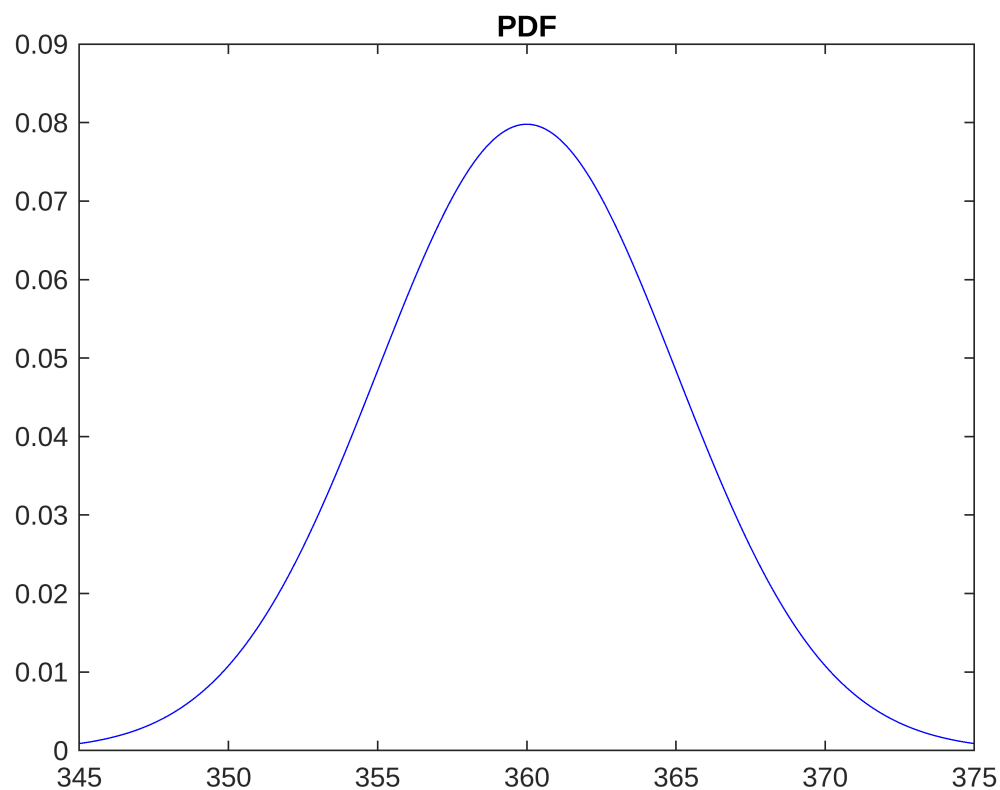
2. O volume de enchimento de uma máquina de enchimento automatizada usada para encher refrigerantes é normalmente distribuído com uma média de 360ml e um desvio padrão de 5ml.

Abaixo temos o gráfico que mostra a PDF do problema:

```
media = 360; % 360ml de média de enchimento normalmente distribuidos
sigma = 5; % 5ml de desvio padrão
x = media - 3*sigma: 0.1: media+3*sigma
```

```
x = 1x301
    345.0000    345.1000    345.2000    345.3000    345.4000    345.5000    345.6000    345.7000 ...
```

```
pdf = normpdf(x, media, sigma);
figure
plot(x, pdf, 'blue')
ylim([0, 0.09])
title("PDF")
```



2.1 Qual a probabilidade de que o volume seja menor que 355ml?

```
volume_prob2_1 = find(x<355);
probabilidade2_1 = trapz(x(volume_prob2_1), pdf(volume_prob2_1))
```

```
probabilidade2_1 = 0.1525
```

Fazendo uma análise estatística, é necessário calcular a integral do gráfico de forma que os limites de integração sejam limitados para valores até $x < 355$ como calculado pelo código acima.

Dessa forma, foi encontrado o resultado sendo uma probabilidade de 15,25%

2.2 Se os refrigerantes com menos de 350ml e mais de 370ml forem eliminados, qual a proporção de refrigerantes eliminados?

```
volume_prob2_2_less = find(x<350);
volume_prob2_2_bigger = find(x>370);
probabilidade2_2 = (trapz(x(volume_prob2_2_less), pdf(volume_prob2_2_less))
+ trapz(x(volume_prob2_2_bigger), pdf(volume_prob2_2_bigger)))

probabilidade2_2 = 0.0407
```

Da mesma forma que calcular o resultado concatenando as áreas que correspondem aos refrigerantes eliminados para os volumes menores que 350ml e maiores que 370ml. Sendo assim, a proporção encontrada é de 4 refrigerantes eliminados para 100 produzidos.

2.3 Determine as especificações de volume de modo que 99% dos refrigerantes sejam aceitos.

```
for xaxis = 360:1:380
    volume_prob2_3 = find(x<xaxis);
    probabilidade2_3 = trapz(x(volume_prob2_3), pdf(volume_prob2_3),2);
    if probabilidade2_3 > 0.99
        X = ['Para uma especificacao de ', num2str(xaxis), ' temos uma prob
de ', num2str(probabilidade2_3), ' serem aceitos'];
        disp(X)
    end
end
```

```
Para uma especificacao de 373 temos uma prob de 0.99371 serem aceitos
Para uma especificacao de 374 temos uma prob de 0.99593 serem aceitos
Para uma especificacao de 375 temos uma prob de 0.99721 serem aceitos
Para uma especificacao de 376 temos uma prob de 0.9973 serem aceitos
Para uma especificacao de 377 temos uma prob de 0.9973 serem aceitos
Para uma especificacao de 378 temos uma prob de 0.9973 serem aceitos
Para uma especificacao de 379 temos uma prob de 0.9973 serem aceitos
Para uma especificacao de 380 temos uma prob de 0.9973 serem aceitos
```

Com o código acima, foi possível encontrar a porcentagem de refrigerantes aceitos para um intervalo de especificações que foi arbitrariamente escolhido de 360ml até 380, em seguida foi mostrado apenas os valores que satisfazem a especificação com 99% de refrigerantes aceitos na produção. Com isso, foi possível determinar que para uma linha de produção para 373ml, temos que 99% dos refrigerantes serão aceitos.

3. Seja uma variável aleatória X com distribuição binomial. Escolha um valor de $n > 10$ e p conforme abaixo.

3.1 Escolha p tal que X possa ser aproximada por uma variável aleatória Y com distribuição normal, plotando os gráficos da distribuição de X e de Y.

Para este problema, será utilizado n de tal forma que $np > 5$ e $n(1 - p) > 5$.

Escolhendo $n = 15$, então $p > \frac{5}{15} = 0.33$, assim podemos escolher $p = 0.4$ e todas inequações são atendidas.

Utilizando a equação abaixo, podemos calcular μ e σ para o problema:

If X is a binomial random variable with parameters p and n ,

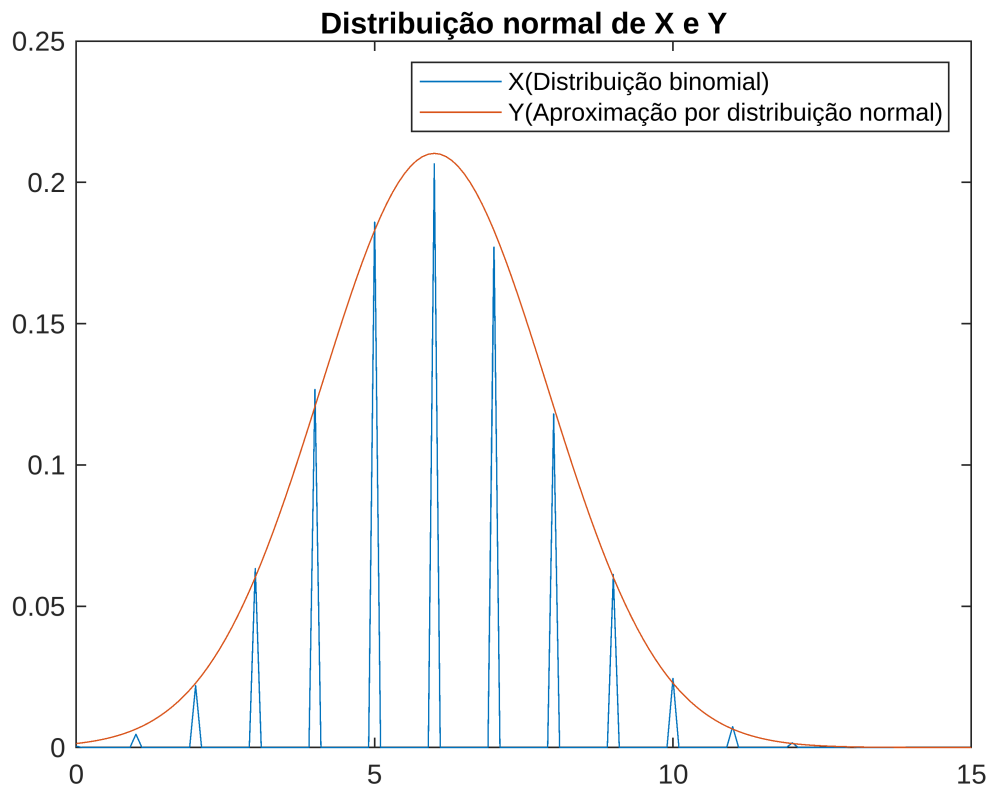
$$\mu = E(X) = np \quad \text{and} \quad \sigma^2 = V(X) = np(1 - p)$$

```
n = 15;
p = 0.4;

axisX = 0:0.1:n;
X = binopdf(axisX,n,p);

mu2 = n*p;
sigma2 = sqrt(n*p*(1-p));
Y = normpdf(axisX,mu2,sigma2);

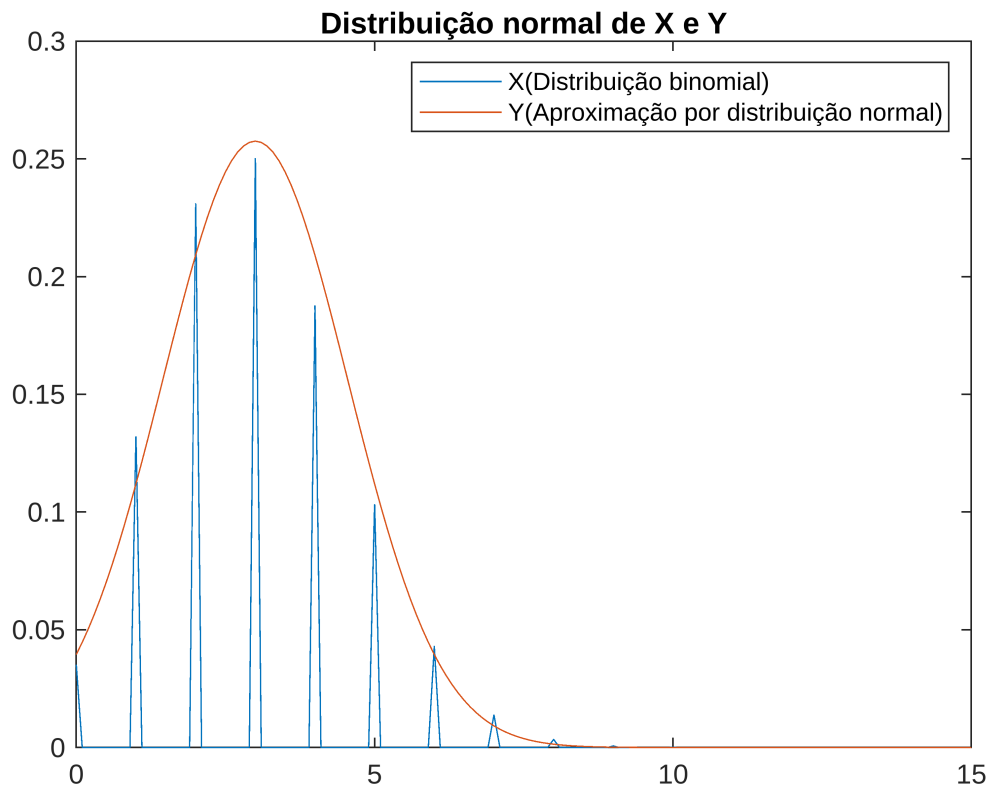
figure
plot(axisX,X)
hold on
plot(axisX,Y)
title("Distribuição normal de X e Y")
legend("X(Distribuição binomial)","Y(Aproximação por distribuição normal)")
```



3.2 Repita para p tal que esta aproximação não seja possível, mostrando isso através dos gráficos de distribuição correspondentes.

Para satisfazer a não possibilidade da aproximação, basta escolhermos p de tal forma que $p < 0.33$, ou seja, se $p = 0.2$ veremos que a aproximação não será possível.

```
p2 = 0.2;
% Criando X
inter = 0:0.1:n;
X2 = binopdf(inter,n,p2);
% Criando Y
mu3 = n*p2;
sigma3 = sqrt(n*p2*(1-p2));
Y2 = normpdf(inter,mu3,sigma3);
figure
plot(inter,X2)
hold on
plot(inter,Y2)
title("Distribuição normal de X e Y")
legend("X(Distribuição binomial)","Y(Aproximação por distribuição normal)")
```



Com o gráfico acima, fica evidente que a aproximação falha em representar perfeitamente a distribuição binomial.

4. Gere uma variável aleatória X com distribuição normal padronizada contendo 500 amostras.

4.1 Faça um histograma de X e sobre ele plote a pdf (0,1)

4.2 Faça um histograma de $Y = X^2$ e sobre ele plote a pdf

4.3 Obtenha $P(-2 \leq X \leq 2)$ e y tal que $P(Y \leq y) = P(-2 \leq X \leq 2)$

4.4 Plote a pdf χ^2_n para $n = 1, 2, \dots, 10$, explique o efeito de n sobre a pdf, e para $n = 10$ compare χ^2_n com $(n, 2n)$.

5. Selecione aleatoriamente 20 amostras de X e construa o intervalo de confiança para a média com $\alpha = 5\%$.

Considere a variância conhecida, obtida da população de 1000 amostras. Repita este procedimento e construa 50 intervalos para a média. Plote os intervalos das médias junto com a média de X e comente os resultados que obteve, comparando os intervalos estimados e verificando se eles contêm a média conhecida da população.

6. Plote o histograma cumulativo das 50 médias calculadas e compare com a cdf teórica da média amostral de X .

7. Repita a atividade 5 fazendo um intervalo de confiança para a variância da variável X gerada com distribuição normal, média μ e variância σ escolhidas. Faça o histograma das 50 variâncias e analise.

8. Um fabricante de equipamentos seleciona aleatoriamente 1000 unidades e verifica que 5 têm defeito.

8.1. Construa um intervalo de confiança de 95% para a proporção de unidades com defeito.

8.2. Há evidências para suportar a afirmação de que a fração de peças com defeito é menor que 1%?