Lista 3

Ítallo Silva - 118110718 | Thiago Nascimento - 118110804 | João Marcelo Junior - 117110448

Questão 1

a)

Sendo as v.a. de X são independentes e cada v.a. segue a distribuição de probabilidade X, temos que as v.a. são amostras aleatórias simples de X, assim sendo, podemos assumir pelo teorema que $E(\overline{\mathrm{X}}) = E(X)$, ou seja, $E(\overline{\mathrm{X}}) = 54$ meses

b)

Uma vez que, as v.a. de X constituem uma A.A.S. e com n = 50, sendo suficientemente grande, podemos assumir pelo TCL que $\overline{\mathrm{X}} \sim N(\mu; \frac{\sigma^2}{n})$, logo temos que, $P(\overline{\mathrm{X}} > 52) = P(Z > \frac{52-54}{\frac{6}{\sqrt{50}}}) = P(Z > \frac{-2}{0,8485}) \simeq P(Z > -2,36)$, sabendo que $P(Z > -2,36) = 1 - P(Z < -2,36)$, pela tabela temos que, $P(Z < -2,36) = 0,0091$, então $P(Z > -2,36) = 1 - 0,0091 = 0,9909$. Sendo assim, pela forte probabilidade de uma bateria durar no mínimo 52 meses a afirmação do fabricante de $\mu = 54$ meses e $\sigma = 6$ meses parece verdadeira.

Questão 2

a)

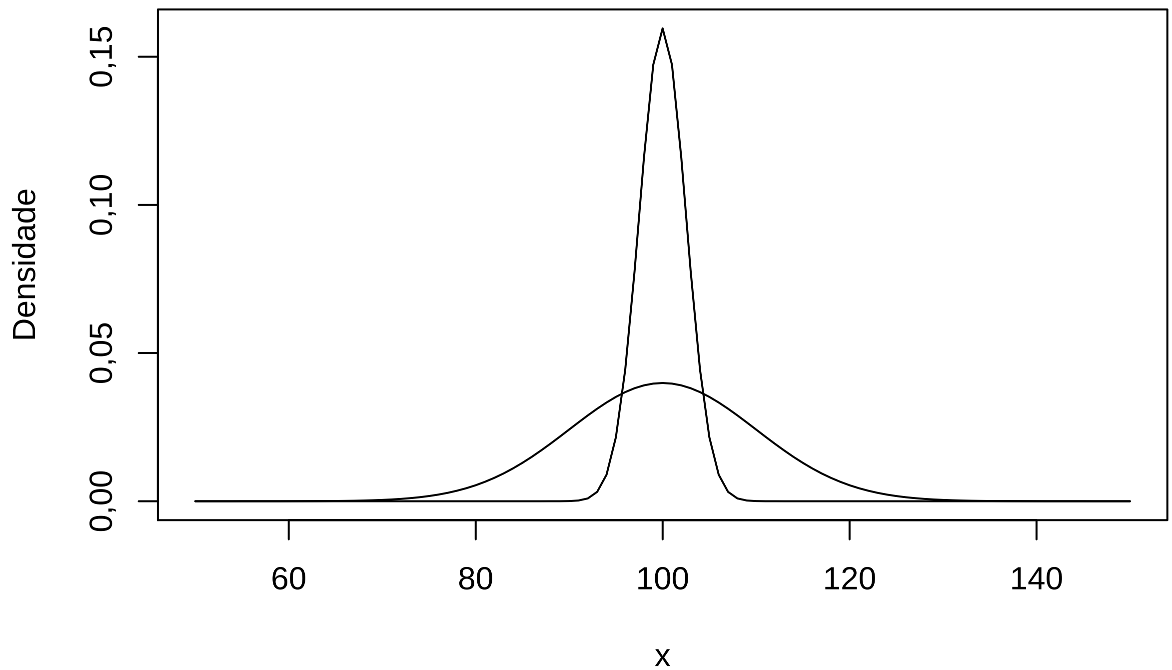
Sabendo que $X \sim N(100; 10^2)$, temos que, $P(90 < X < 110) = P(\frac{90-100}{10} < Z < \frac{110-100}{10}) = P(-1 < Z < 1)$, sendo $Z \sim N(0;1)$, temos que $P(-1 < Z < 1) = P(Z < 1) - P(Z < -1)$, logo $P(-1 < Z < 1) = 0,8413 - 0,1587 = 0,6826$

b)

Sabendo que $X \sim N(100; 10^2)$, então $\overline{\mathrm{X}} \sim (\mu; \frac{\sigma^2}{n})$, temos que $\overline{\mathrm{X}} \sim (100; \frac{10^2}{16})$, logo $P(90 < \overline{\mathrm{X}} < 110) = P(\frac{90-100}{\frac{10}{\sqrt{16}}} < Z < \frac{110-100}{\frac{10}{\sqrt{16}}}) = P(-4 < Z < 4)$, sendo $Z \sim N(0,1)$, temos que $P(-4 < Z < 4) = P(Z < 4) - P(Z < -4)$, logo $P(-4 < Z < 4) \simeq 1,0000 - 0,0000 \simeq 1,000$

c)

```
f <- function(x) {dnorm(x, mean = 100, sd = 10)}
g <- function(x) {dnorm(x, mean = 100, sd = 2.5)}

curve(g, from=50, to=150, ylab = "Densidade")
curve(f, from=50, to=150, add=TRUE)
```



d)

Sabendo que $\overline{\mathrm{X}} \sim (\mu; \frac{\sigma^2}{n})$, temos que $\overline{\mathrm{X}} \sim (100; \frac{10^2}{n})$, sendo assim, $P(90 < \overline{\mathrm{X}} < 110) = P(\frac{90-100}{\frac{10}{\sqrt{n}}} < Z < \frac{110-100}{\frac{10}{\sqrt{n}}}) = P(-\sqrt{n} < Z < \sqrt{n})$, e $P(-\sqrt{n} < Z < \sqrt{n}) = P(Z < \sqrt{n}) - P(Z < -\sqrt{n})$, e $P(Z < \sqrt{n}) - P(Z < -\sqrt{n}) = P(Z < \sqrt{n}) - [1 - P(Z < \sqrt{n})] = 2P(Z < \sqrt{n}) - 1$, logo $2P(Z < \sqrt{n}) - 1 = 0,95 \rightarrow P(Z < \sqrt{n}) = \frac{1,95}{2} = 0,975$, sendo $Z \sim N(0;1)$, pela tabela $P(Z < 1,96) = 0,975$, então $\sqrt{n} = 1,96 \rightarrow n = 1,96^2 = 3,842$

Questão 3

a)

Sendo $X \sim N(\mu; 10^2)$, temos que $P(x < 500) = P(\frac{x-\mu}{10} < \frac{500-\mu}{10}) = P(z < \frac{500-\mu}{10}))$, sendo $Z \sim N(0;1)$, então $P(z < \frac{500-\mu}{10}) = 0,1$, pela tabela temos que $P(z < -1,28) \simeq 0,1$, logo $\frac{500-\mu}{10} = -1,28 \rightarrow -\mu = -12,8 - 500 \rightarrow \mu = 512,8g$

b)

Sabendo que $X \sim N(512,8; 10^2)$, então $\overline{\mathrm{X}} \sim N(512,8; \frac{10^2}{4})$, sendo assim, $P(\overline{\mathrm{X}} \leq 500) = P(Z \leq \frac{500-512,8}{\frac{10}{\sqrt{4}}}) = P(Z < -2,56) = 0,0052$

Questão 4

a)

Se a máquina está regulada: $\overline{\mathrm{X}} \sim N(512,8; \frac{10^2}{4})$

Parada desnecessária ocorre quando $\overline{\mathrm{X}} < 495$ ou $\overline{\mathrm{X}} > 520$, portanto:

$P(\overline{\mathrm{X}} < 495 \text{ ou } \overline{\mathrm{X}} > 520) = P(Z \leq \frac{495-512,8}{\frac{10}{\sqrt{4}}}) + (1 - P(Z \leq \frac{520-512,8}{\frac{10}{\sqrt{4}}})) = P(Z < -3,56) + (1 - P(Z < 1,44))$ $= 0,002 + 0,0749 = 0,0751 = 7,51\%$

b)

Se o peso médio desregulou-se para 500g: $\overline{\mathrm{X}} \sim N(500; \frac{10^2}{4})$.

P(continuar fora dos padrões) = $P(495 \leq \overline{\mathrm{X}} \leq 520) = P(Z \leq \frac{520-500}{\frac{10}{\sqrt{4}}})$ - $P(Z \leq \frac{495-500}{\frac{10}{\sqrt{4}}})) = P(Z < 4) - P(Z < -1) = 1 - 0,1587 = 0,8413 = 84,13\%$

Questão 5

a)

$\overline{\mathrm{X}} \sim N(70; \frac{10^2}{7})$

$P(\sum_{i=1}^{7} X_i > 500) = P(\overline{\mathrm{X}} > \frac{500}{7}) = 1 - P(Z < \frac{\frac{500}{7} - 70}{\frac{10}{\sqrt{7}}}) = 1 - P(Z < 0,38) = 1 - 0,6480 = 0,352 = 35,2\%$

b)

$\overline{\mathrm{X}} \sim N(70; \frac{10^2}{6})$

$P(\sum_{i=1}^{6} X_i > 500) = P(\overline{\mathrm{X}} > \frac{500}{6}) = 1 - P(Z < \frac{\frac{500}{6} - 70}{\frac{10}{\sqrt{6}}}) = 1 - P(Z < 3,27) = 1 - 0,9995 = 0,0005 = 0,05\%$

Questão 6

Vamos considerar uma população em que a proporção de elementos portadores de certa característica é p. Logo, podemos definir uma v.a. X, da seguinte maneira:

X = 1, se o indíviduo for portador da característica, 0, se o indívdiuo não for portador da característica

Logo,

$\mu = E(X) = p$, $\sigma^2 = Var(X) = p(1-p)$

Aplicando o Teorema Central do Limite:

Quando X1, . . . , Xn é uma amostra aleatória simples retirada de uma população $X \sim Ber(p)$ cuja média é, portanto, $E(X) = \mu = p$ e variância $Var(X) = \sigma^2 = p(1-p)$ com 0 < p < 1; a distribuição amostral da proporção de sucessos $\overline{P}$ aproxima-se para n grande de uma distribuição normal:

$\overline{P} \sim N(p, \frac{p(1-p)}{n}$, para n suficientemente grande.

Questão 7

Seja p a proporção de mulheres, então $p = 0,30$.

Pelo TLC, sabemos que $\hat{p} \sim N\big(p; \frac{p(1-p)}{n}\big)$, substituindo p e n, temos que $\hat{p} \sim N(0,3; 0,021)$.

A probabilidade que buscamos é dada por $P(|\hat{p} - p| < 0,01) = P(-0,01 < \hat{p} - p < 0,01)$. Somando p na igualdade temos $P(|\hat{p} - p| < 0,01) = P(0,29 < \hat{p} < 0,31)$.

Padronizando, temos: $P(0,29 < \hat{p} < 0,31) = P(\frac{0,29-0,30}{\sqrt{0,021}} < Z < \frac{0,31-0,30}{\sqrt{0,021}}) = P(-0,07 < Z < 0,07) = P(Z < 0,07) - P(Z < -0,07) = 0,5279 - 0,4721 = 0,056$.

Assim $P(|\hat{p} - p| < 0,01) = 0,056$.

Questão 8

a)

Temos que $S_n \sim Bin(8; 0,20)$, assim:

```
dist <- list('n' = as.character(seq(0,8)),
             '$P(S_n = n)$' = round(dbinom(seq(0,8), 8, 0.2), 3)) %>%
  as_tibble() %>%
  t() %>%
  as_tibble()

rownames(dist) <- c("n", "$P(S_n = n)$")

dist %>% knitr::kable(digits = 3, col.names = rep("", 9))
```

n	0	1	2	3	4	5	6	7	8
$P(S_n = n)$	0,168	0,336	0,294	0,147	0,046	0,009	0,001	0,000	0,000

b)

Temos que 25% de peças defeituosas corresponde a $25\% \cdot 8 = 2$ peças. Assim, desejamos obter $P(S_n \leq 2)$. Usando R:

```
p <- pbinom(2, 8, 0.20)
```

Assim, $P(S_n \leq 2) =$ 0,7969.

c)

Pelo TCL, temos que $S_n \sim N(np; npq)$, onde $n = 8$, $p = 0,20$ e $q = 0,80$, assim $S_n \sim N(1,6; 1,28)$. Agora, calculemos $P(S_n \leq 2)$, correspondente a no máximo 25% de peças com defeito:

$P(S_n \leq 2) \rightarrow P\big(Z \leq \frac{2-1,6}{\sqrt{1,28}}\big) = P(Z \leq 0,35)$.

Usando a tabela temos que $P(Z \leq 0,35) = 0,6368$. Calculando no R, temos:

```r
p <- pnorm(0.35)
```

Logo $P(Z \leq 0,35) = 0{,}6368$. A probabilidade foi bem diferente, isso se deve ao fato do n ser pequeno e $n \cdot p < 5$, assim fazendo a aproximação não ser muito boa.

d)

Considerando um sorteio de 30 peças. Temos que $25\% \cdot 30 = 7.5$.

Usando a distribuição binomial (item b), temos:

```r
pb <- pbinom(7, 30, 0.20) # Como a variável é discreta 7, pode ser usando ao invés de 7,5
```

Usando a aproximação pela Normal (item c), temos:

$S_n \sim N(np; npq) \rightarrow S_n \sim N(6; 4,8)$, assim calculando no R:

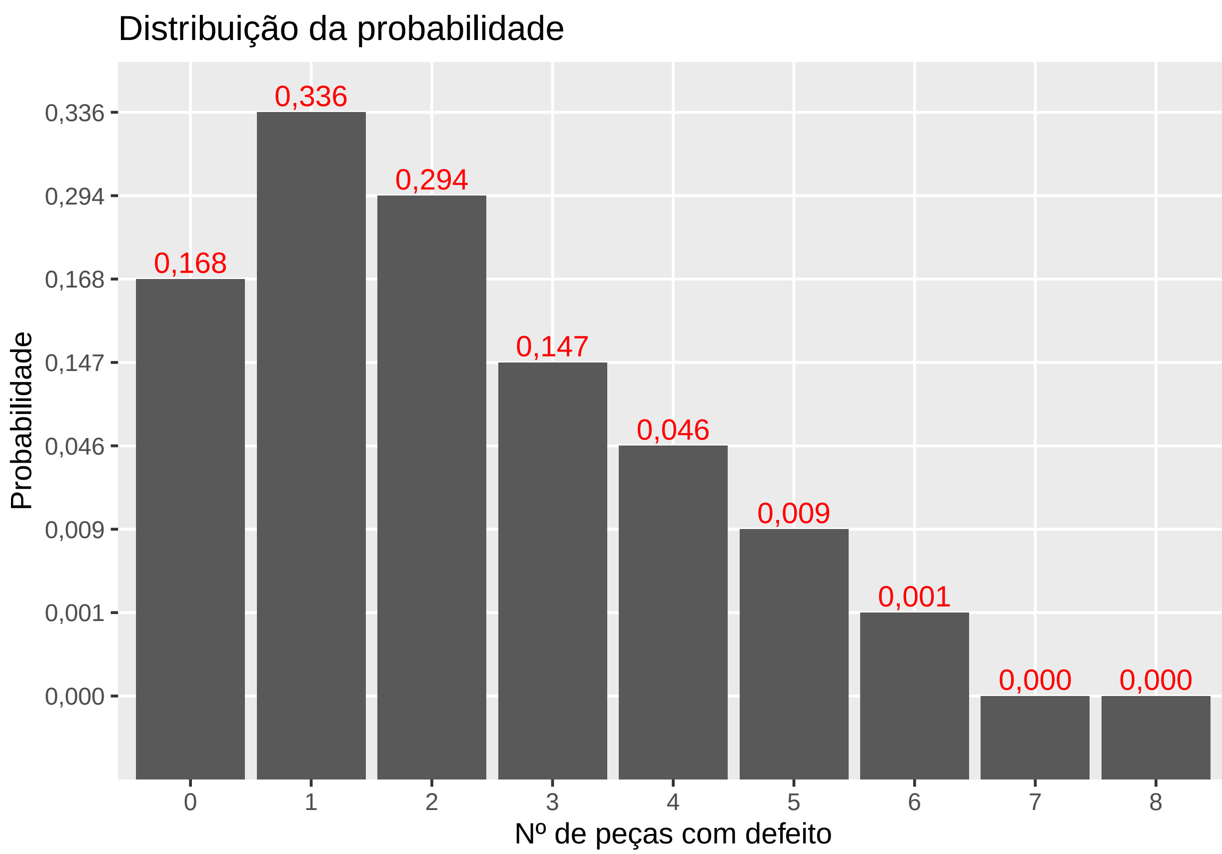
```r
pn <- pnorm(7.5, mean = 6, sd = sqrt(4.8))
```

Pela binomial temos $P(S_n \leq 7.5) = 0{,}7608$ e pela normal temos $P(S_n \leq 7.5) = 0{,}7532$. Com $n = 30$, temos uma aproximação bem mais próxima do que as calculadas com $n = 5$, nos itens b e c.

e)

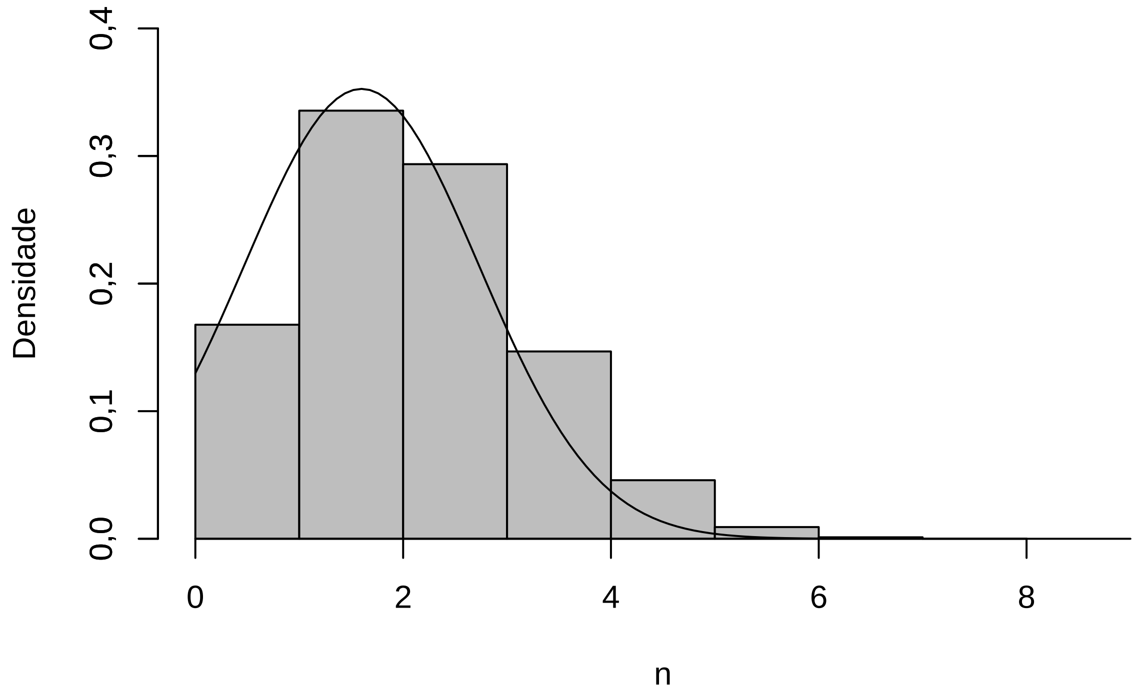
Representação do item a

```r
dist %>% t() %>% as_tibble() %>%
  ggplot(aes(x = n, y = `$P(S_n = n)$`)) +
  geom_bar(stat = "identity", position = 'dodge') +
  labs(title = "Distribuição da probabilidade",
       x = "Nº de peças com defeito",
       y = "Probabilidade") +
  geom_text(mapping = aes(x = n, y = `$P(S_n = n)$`, label = `$P(S_n = n)$`),
            color = 'red',
            nudge_y = 0.2)
```



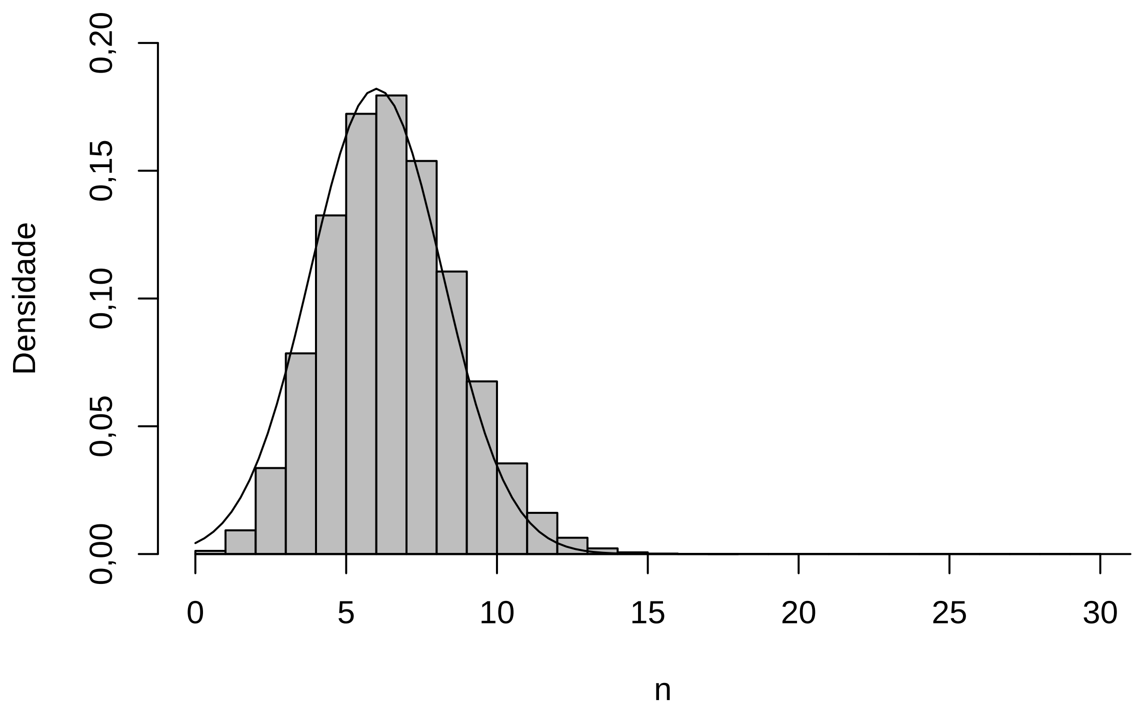
Representação do item b e c

```
f <- function(x) { dnorm(x, mean = 1.6, sd = sqrt(1.28)) }

barplot(dbinom(seq(0, 8), 8, 0.2), ylim = c(0, .4), space = 0, xlab = "n", ylab = "Densidade")
curve(f, from = 0, to = 8, add = TRUE)
axis(1)
axis(2)
```



Representação do item d

```
f <- function(x) { dnorm(x, mean = 6, sd = sqrt(4.8)) }

barplot(dbinom(seq(0, 30), 30, 0.2), ylim = c(0, .2), axes = FALSE, space = 0, xlab = "n", ylab = "Dens
curve(f, from = 0, to = 30, add = TRUE)
axis(1)
axis(2)
```



Questão 9

A parada é desnecessária caso o processo ainda esteja dentro da margem de 10% itens defeituosos. No pior caso, sendo o próprio 10%.

Sendo assim, o número Y de peças defeituosas no sorteio de 20 peças tem distribuição $Y \sim Bin(20; 0,1)$ e a proporção de peças com defeito $\hat{p}$ tem distribuição aproximadamente $N(0,1; \frac{0,1 \cdot 0,9}{20}) = N(0,1; 0,0045)$.

Queremos $P(\hat{p} > 0,15) \rightarrow P(Z > 0,75) = 1 - P(Z < 0,75) = 1 - 0,7734 = 0,2266$. Assim a probabilidade de uma parada desnecessária é $22,66\%$.

Questão 10

a)

Considerando a questão anterior, temos que a proporção $\hat{p}$ de peças em uma caixa de 100 unidades tem distribuição aproxida por $N(0,1; 0,0009)$. E queremos $P(\hat{p} > 0,1)$, assim:

$P(\hat{p} > 0,1) = P\left(Z > \frac{0,1-0,1}{\sqrt{0,0009}}\right) = P(Z > 0) = 50\%$.

b)

Queremos $P(\hat{p} = 0)$, sendo assim não é recomendado utilizar a distribuição normal já que queremos uma probabilidade pontual, mas podemos utilizar a distribuição binomial $\hat{p} \sim Bin(100; 0.1)$. Sendo assim, $P(\hat{p} = 0) \approx 2,65 \times 10^{-5} \approx 0$.

c)

```
f <- function(x) { dnorm(x, mean = 10, sd = 3) }

barplot(dbinom(seq(0, 100), 100, 0.1), space = 0, ylim = c(0, .15), axes = FALSE, xlab = "n", ylab = "De
curve(f, add = TRUE)
axis(1)
axis(2)
```