

```
#libraries
library(tibble)
```

Sabe-se que nos municípios de uma certa região turística, 40% [destes municípios] mais que duplicam a população nos meses de verão.

Suponha que o consumo de água por cada turista alojado, por dia, pode ser descrito através de uma variável aleatória $[X]$ com distribuição Normal, de média 0,50 m³ e desvio padrão 0,05 m³.

A função de densidade de probabilidade de uma variável aleatória $X \sim N(\mu, \sigma^2)$ com distribuição Normal é dada por:

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (1)$$

Logo, a função de distribuição [cumulativa] de probabilidade de $X \sim N(0.5, 0.05^2)$ é dada por:

$$F(x) = P\{X \leq x\} = \int_{-\infty}^x f(t|\mu = 0.5, \sigma^2 = 0.05^2) dt \quad (2)$$

$$= \frac{1}{0.05\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{(t-0.5)^2}{2(0.05)^2}\right) dt \quad (3)$$

$$= 20(2\pi)^{-1/2} \int_{-\infty}^x \exp\left(-\frac{(t-0.5)^2}{0.005}\right) dt \quad (4)$$

, sendo x o consumo de água por turista alojado, por dia.

a)

Qual a probabilidade do consumo, por turista e por dia, ser inferior a 0,437 m³?

Como a função é contínua,

$$P\{X < x\} = P\{X \leq x\} = F(x). \quad (5)$$

Conforme a Eq. (4):

$$P\{X < 0.437\} = 20(2\pi)^{-1/2} \int_{-\infty}^{0.437} \exp\left(-\frac{(t-0.5)^2}{0.005}\right) dt \quad (6)$$

```
intF <- function(x) {
  exp(-((x - 0.5)^2) / 0.005)
}
print(20 * (2 * pi)^(-1/2) * integrate(intF, lower = -Inf, upper = 0.437)$
      value)
```

```
## [1] 0.1038347
```

```
print(pnorm(0.437, mean = 0.5, sd = 0.05, lower.tail = TRUE))
```

```
## [1] 0.1038347
```

b)

Calcule o maior consumo dos 25% menores.

Conforme a Eq. 6,

$$P\{X \leq x\} = 0.25 \quad (7)$$

$$20(2\pi)^{-1/2} \int_{-\infty}^x \exp\left(-\frac{(t-0.5)^2}{0.005}\right) dt = 0.25 \quad (8)$$

$$\int_{-\infty}^x \exp\left(-\frac{(t-0.5)^2}{0.005}\right) dt = 0.0125(2\pi)^{1/2} \quad (9)$$

```
rightVal <- 0.0125 * (2 * pi)^(1/2)
print(pracma::bisect(function(x) integrate(intF, lower = -Inf, upper = x)$
  value - rightVal, 0, 1)$root)
```

```
## [1] 0.4662755
```

```
print(qnorm(0.25, mean = 0.5, sd = 0.05, lower.tail = TRUE))
```

[1] 0.4662755

c)

Numa certa unidade de turismo rural casa, a capacidade é de 20 hóspedes. Considerando um momento em que a unidade está totalmente cheia, qual a probabilidade do consumo de água diário aumentar em pelo menos 10,75 m³ (quando comparado com a situação em que não existem hóspedes)?

Se X_1, X_2, \dots, X_n forem variáveis normais independentes com médias $\mu_1, \mu_2, \dots, \mu_n$ e variâncias $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$,

$$T = \sum_{i=1}^n a_i X_i \sim N \left(\sum_{i=1}^n a_i \mu_i, \sqrt{\sum_{i=1}^n a_i^2 \sigma_i^2} \right) \quad (10)$$

Como $a_i = 1 \wedge \mu_i = \mu = 0.5 \wedge \sigma_n^2 = \sigma^2 = 0.05^2$,

$$T = 20X \sim N(20\mu, \sqrt{20\sigma^2}) \quad (11)$$

$$\sim N(20 \cdot 0.5, \sqrt{20 \cdot 0.05^2}) \quad (12)$$

$$\sim N(10, 0.05) \quad (13)$$

Conforme a Eq. 1,

$$P\{T > 10.75\} = 1 - P\{T \leq 10.75\} \quad (14)$$

$$= 1 - \int_{-\infty}^{10.75} f(t|\mu = 20 \cdot 0.5, \sigma = \sqrt{0.05}) dt \quad (15)$$

$$= 1 - \frac{1}{\sqrt{0.05}\sqrt{2\pi}} \int_{-\infty}^{10.75} \exp \left(-\frac{(t-10)^2}{2 \cdot (\sqrt{0.05})^2} \right) dt \quad (16)$$

$$= 1 - (0.1\pi)^{-1/2} \int_{-\infty}^{10.75} \exp \left(-\frac{(t-10)^2}{0.1} \right) dt \quad (17)$$

```
intF <- function(x) {
  exp(-((x - 10)^2) / 0.1)
}
print(1 - (0.1 * pi)^(-1/2) * integrate(intF, lower = -Inf, upper = 10.75)
      $value)
```

```
## [1] 0.0003981151
```

```
print(1 - pnorm(10.75, mean = 20*0.5, sd = sqrt(20)*0.05, lower.tail =
      TRUE))
```

```
## [1] 0.0003981151
```

d)

Em relação à casa de turismo rural da alínea anterior:

i

Simule *uma* observação do acréscimo de consumo de água gerado pelos 20 hóspedes dessa casa de turismo rural.

```
S <- rnorm(20, mean = 0.5, sd = 0.05)
tibble(sum(S), mean(S), min(S), max(S))
```

```
## # A tibble: 1 x 4
##   `sum(S)` `mean(S)` `min(S)` `max(S)`
##   <dbl>    <dbl>    <dbl>    <dbl>
## 1      9.78      0.489      0.409      0.559
```

O consumo simulado pelos 20 hóspedes variam entre 0.409 m^3 e 0.559 m^3 , e os 20 hóspedes coletivamente consumiram 9.784 m^3 .

ii

Simule 1000 observações nas mesmas condições, guardando *apenas* a soma para cada uma delas

```
times <- 1000
S <- colSums(matrix(rnorm(20*times, mean = 0.5, sd = 0.05), ncol = times))
tibble(S)
```

```
## # A tibble: 1,000 x 1
##       S
##   <dbl>
## 1  9.83
## 2  9.76
## 3 10.1
## 4 10.4
## 5 10.2
## 6 10.2
## 7  9.76
## 8  9.84
## 9 10.2
## 10  9.97
## # ... with 990 more rows
```

[...] com base nessa simulação, estime a probabilidade que calculou de forma exata em c).

```
print(sum(S > 10.75) / times)
```

```
## [1] 0
```

```
print(1 - pnorm(10.75, mean = 20*0.5, sd = sqrt(20)*0.05, lower.tail =
TRUE))
```

```
## [1] 0.0003981151
```

Sendo que a probabilidade exata para o consumo é tão pequena, a simulação não é muito precisa com 1000 observações. Para demonstrar uma simulação mais precisa, vamos aumentar o número de observações para 1 000 000.

```
times <- 1000000
S <- colSums(matrix(rnorm(20*times, mean = 0.5, sd = 0.05), ncol = times))
print(sum(S > 10.75) / times)
```

```
## [1] 0.000405
```

```
print(1 - pnorm(10.75, mean = 20*0.5, sd = sqrt(20)*0.05, lower.tail =  
      TRUE))
```

```
## [1] 0.0003981151
```

Agora torna-se mais evidente certidão da probabilidade exata.