Lista de Exercicios IPE#4b

Taiguara Melo Tupinambas

Entrega: 06 de maio de 2017

Exercício 1

O valor esperado de X_i, $E[X_i] = \frac{1}{2} - \frac{1}{2} = 0$

Logo, o valor esperado de Y_N é temos que:

$$Y_N = \sum_{i=1}^{N} \frac{X_i}{\sqrt{N}} \tag{1}$$

$$E[Y_N] = E[\sum_{i=1}^{N} \frac{X_i}{\sqrt{N}}] = \frac{1}{\sqrt{(N)}} N E[X] \tag{2}$$

$$E[Y_N] = 0 \tag{3}$$

A variância de X_i, $var(X_I) = E[X^2] - E^2[X] = E[X^2]$ e como $E[X^2] = \frac{1}{2} + \frac{1}{2} = 1$, temos que a variância de Y_N é dada por:

$$var(Y_N) = var(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} X_i) = \frac{1}{N} \sum_{i=1}^{N} var(X_i) = var(X) \tag{4}$$

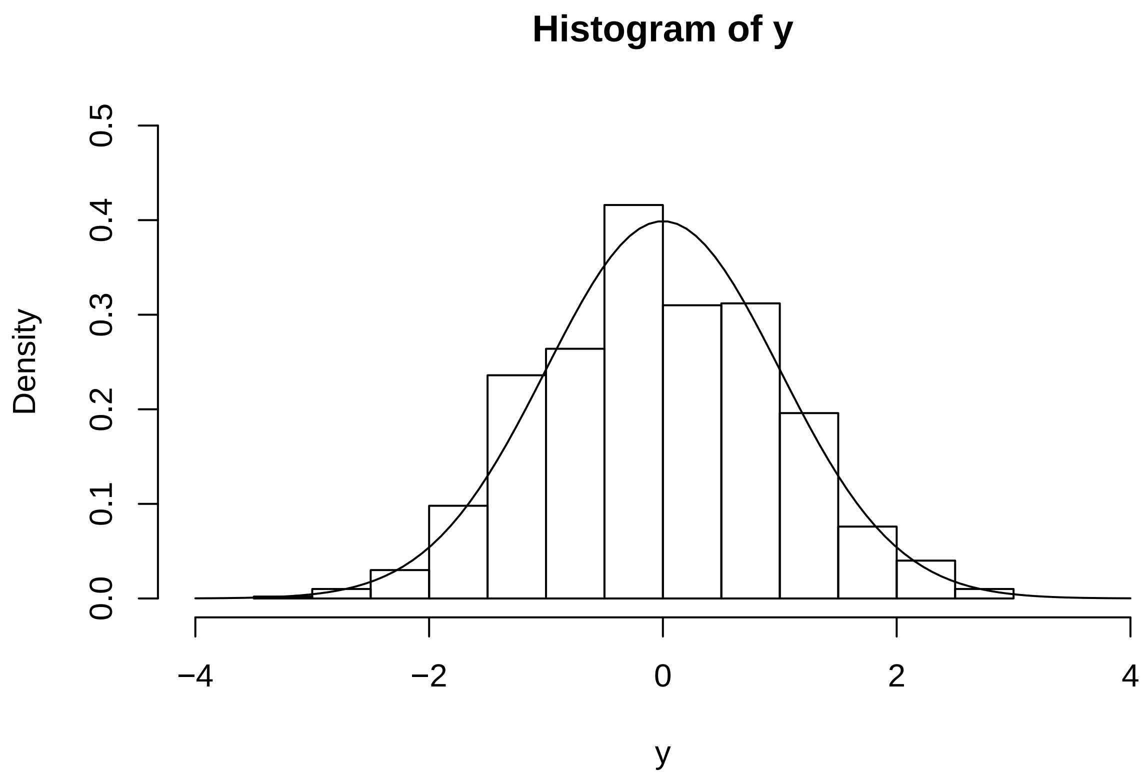
$$var(Y_N) = 1 \tag{5}$$

Logo, pelo teorema do limite central, $Y_N \sim N(0,1)$. Realizando a simulação, os resultados encontrados foram coerentes.

```r
set.seed(1)

N<-1000
n<-1000
y<-matrix(0,N)

for (i in 1:N) {
        for (j in 1:n) {
                if (runif(1)<0.5) {
                        x<-1
                } else {
                        x<--1
                }
                y[i]<-y[i]+x
        }
}

y<-y/sqrt(n)

cat("Média = ", mean(y), "Varância =", var(y))
```

```
## Média =  -0.02340085 Varância = 1.007116
```

```r
x<-seq(-4,4,length=100)
distx<-dnorm(x)

hist(y,freq = F,xlim=c(-4,4),ylim=c(0,0.5))
lines(x,distx)
```



Exercício 2

Para uma sequência de VAs independentes de Bernoulli dadas por:

$$p_X(X_i) = \begin{cases} p, & X_i = 1 \\ 1-p, & X_i = 0 \end{cases}$$

Desta forma, o estimador de p pode ser definido como $p = E[Y_N]$, em que $Y_N = \frac{1}{N}\sum_{i=1}^{N} X_i$ e a variância de Y_N é igual a $\frac{p(1-p)}{N}$.

Pela desigualdade de Chebyshev, temos que:

$$P(|\hat{Y}_N - E[Y_N]| > \epsilon) \leq \frac{var(Y_N)}{\epsilon^2} \tag{6}$$

$$P(|\hat{p} - p| > \epsilon) \leq \frac{p(1-p)}{N\epsilon^2} \tag{7}$$

A definição de um valor de N, tal que a probabilidade do valor estimado para p ser, em módulo, menor do que 0,01, de 95%, é o mesmo que dizer:

$$P(|\hat{p} - p| < 0.01) \geq 95\% \tag{8}$$
$$\tag{9}$$
$$P(|\hat{p} - p| > 0.01) \leq 100\% - 95\% = 5\% \tag{10}$$
$$\tag{11}$$

Ou seja, a última expressão seria a desigualdade de Chebyshev, com $\epsilon = 0.01$, e $\frac{var(Y_N)}{\epsilon^2} = \frac{p(1-p)}{0.01^2 N} = 0.05$

Logo o valor de N, que garante que a estimação de p tenha erro de no máximo 0.01, a uma probabilidade de, pelo menos, 95% é:

$$\frac{var(Y_N)}{\epsilon^2} = 0.05 \tag{12}$$
$$\frac{p(1-p)}{0.01^2 N} = 0.05 \tag{13}$$
$$N = p(1-p)2 * 10^6 \tag{14}$$

Algumas simulaçÕes foram feitas para valroes de p iguais a 0.1 (N=18000), 0.5 (N=50000) e 0.75 (N=37500).

```
set.seed(1)

p<-c(0.1,0.5,0.75)
N<-p*(1-p)*200000
n<-100
y<-matrix(0,length(N),n)

for (k in 1:n) {
        for (i in 1:length(N)) {
                for (j in 1:N[i]) {
                        if (runif(1)<p[i]) {
                                x<-1
                        } else {
                                x<-0
                        }
                        y[i,k]<-y[i,k]+x
                }
                y[i,k]<-y[i,k]/N[i]
        }
}


erro<-abs(p-y)
erro<-erro<0.01

prob_erro<-matrix(0,length(N),n)

for (i in 1:length(N)) {
        prob_erro[i]<-sum(erro[i,])
}
```

```r
cat("Frequência de erros menores que 0.01 para p=0.1", prob_erro[1])
```

```
## Frequência de erros menores que 0.01 para p=0.1 100
```

```r
cat("\nFrequência de erros menores que 0.01 para p=0.5", prob_erro[2])
```

```
##
## Frequência de erros menores que 0.01 para p=0.5 100
```

```r
cat("\nFrequência de erros menores que 0.01 para p=0.75", prob_erro[3])
```

```
##
## Frequência de erros menores que 0.01 para p=0.75 100
```

Exercício 3

Pelo teorema do limite central, a VA $Y_N = \frac{1}{N}\sum_{i=1}^{N} X_i$ terá uma distribuição aproximadamente normal, com média iguala $\mu = E[Y_N] = 0$ e desvio padrão igual a $\sigma = \sqrt{var(Y_N)} = \sqrt{\frac{p(1-p)}{N}}$. Sabemos, no entando, que

Para a distribuição normal, 95% dos valores estarão dentro do intervalo $\pm 1.96\sigma$. Para que, com 95% de probabilidade, o erro absoluto seja menor do que 0.01, temos que:

$$1.96\sigma = 0.01 \tag{15}$$

$$\sigma = \sqrt{var(Y_N)} = \frac{0.01}{1.96} \tag{16}$$

$$\sqrt{\frac{p(1-p)}{N}} = \frac{0.01}{1.96} \tag{17}$$

$$\frac{p(1-p)}{N} = (\frac{0.01}{1.96})^2 \tag{18}$$

$$N = p(1-p)(\frac{1.96}{0.01})^2 \tag{19}$$

Algumas simulaçÕes foram feitas para valroes de p iguais a 0.1 (N=3458), 0.5 (N=9604) e 0.75 (N=7203).

```r
set.seed(1)

p<-c(0.1,0.5,0.75)
N<-ceiling(p*(1-p)*(1.96/0.01)^2)
n<-1000
y<-matrix(0,length(N),n)

for (k in 1:n) {
        for (i in 1:length(N)) {
                for (j in 1:N[i]) {
                        if (runif(1)<p[i]) {
                                x<-1
                        } else {
                                x<-0
                        }
                        y[i,k]<-y[i,k]+x
                }
```

```
                y[i,k]<-y[i,k]/N[i]
        }
}


erro<-abs(p-y)
erro<-erro<0.01

prob_erro<-matrix(0,length(N),n)

for (i in 1:length(N)) {
        prob_erro[i]<-sum(erro[i,])/length(erro[1,])
}

cat("Frequência de erros menores que 0.01 para p=0.1", prob_erro[1])
```

```
## Frequência de erros menores que 0.01 para p=0.1 0.948
```

```
cat("\nFrequência de erros menores que 0.01 para p=0.5", prob_erro[2])
```

```
##
## Frequência de erros menores que 0.01 para p=0.5 0.953
```

```
cat("\nFrequência de erros menores que 0.01 para p=0.75", prob_erro[3])
```

```
##
## Frequência de erros menores que 0.01 para p=0.75 0.949
```

Como a frequência relativa de ocorrência de erros absolutos maiores que 0.01 foi próxima do 95%, o resultado foi conforme esperado.

Interessante notar que o cálculo de N por Chebyshev é muito mais conservador do que pelo TLC.

Exercício 4

15.5

$$X_N = X_{N-1} + U_N \tag{20}$$
$$X_N = U_1 + U_2 + ... + U_N \tag{21}$$
$$X_N = \sum_{i=1}^{N} U_i \tag{22}$$
$$\tag{23}$$

Para avaliar a convergência, podemos analisar a desigualdade de Chebyshev:

$$P[|\bar{X}_N - E_X[X]| > \epsilon] \leq \frac{var(\bar{X}_N)}{\epsilon^2} \tag{24}$$

Mas, como:

$$var(\bar{X}_N) = \sum_{i=1}^{N} var(U_i) = \sum_{i=1}^{N} \frac{(1+1)^2}{12} = \sum_{i=1}^{N} \frac{1}{3} \tag{25}$$

$$\tag{26}$$

$$\lim_{N \to \infty} var(\bar{X}_N) \to \infty \tag{27}$$

X_N não irá convergir.

15.6

Considerando que as variáveis X_i para $i = 1, 2.., N$ são independentes e idênticamente distribuídas, temos que:

$$X_N = \frac{1}{N} \sum_{i=1}^{N} X_i^2 \tag{28}$$

$$E_X[X_N^2] = \frac{1}{N} \sum_{i=1}^{N} E_X[X^2] = E_X[X^2] \tag{29}$$

$$\tag{30}$$

$$var(X_N^2) = E_X[X^4] - E_X^2[X^2] \tag{31}$$

Considerando que $E_X[X_N^2]$ existe, para que o estimador convirja, é necessário que $var(X_N^2) < \infty$

15.7

$$X_N = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} X_i \tag{32}$$

$$\tag{33}$$

$$var(X_N) = var(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} X_i) = \frac{1}{N} var(\sum_{i=1}^{N} X_i) = var(X) \tag{34}$$

Como $var(X) \neq 0$, para $N \to \infty$, a variável aleatória não irá convergir

15.8

Temos que:

$$Y_N = \begin{cases} \frac{X_N}{\sqrt{N}} + 1, \; se \; \frac{X_N}{\sqrt{N}} > 0.1 \\ \frac{X_N}{\sqrt{N}}, \; se \; \frac{X_N}{\sqrt{N}} \leq 0.1 \end{cases}$$

Logo:

$$P[|Y_N|>\epsilon]=P[|\frac{X_N}{\sqrt{N}}+1|>\epsilon]P[\frac{X_N}{\sqrt{N}}>0.1]+P[|\frac{X_N}{\sqrt{N}}|>\epsilon]P[\frac{X_N}{\sqrt{N}}\leq 0.1] \tag{35}$$

$$P[|Y_N|>\epsilon]=(1-P[-\epsilon<\frac{X_N}{\sqrt{N}}+1<\epsilon])P[\frac{X_N}{\sqrt{N}}>0.1]+(1+P[-\epsilon<\frac{X_N}{\sqrt{N}}<\epsilon])P[\frac{X_N}{\sqrt{N}}\leq 0.1] \tag{36}$$

(37)

Sabendo que:

$$\frac{X_N}{\sqrt{N}}\sim N(0,\frac{1}{\sqrt{N}}) \tag{38}$$

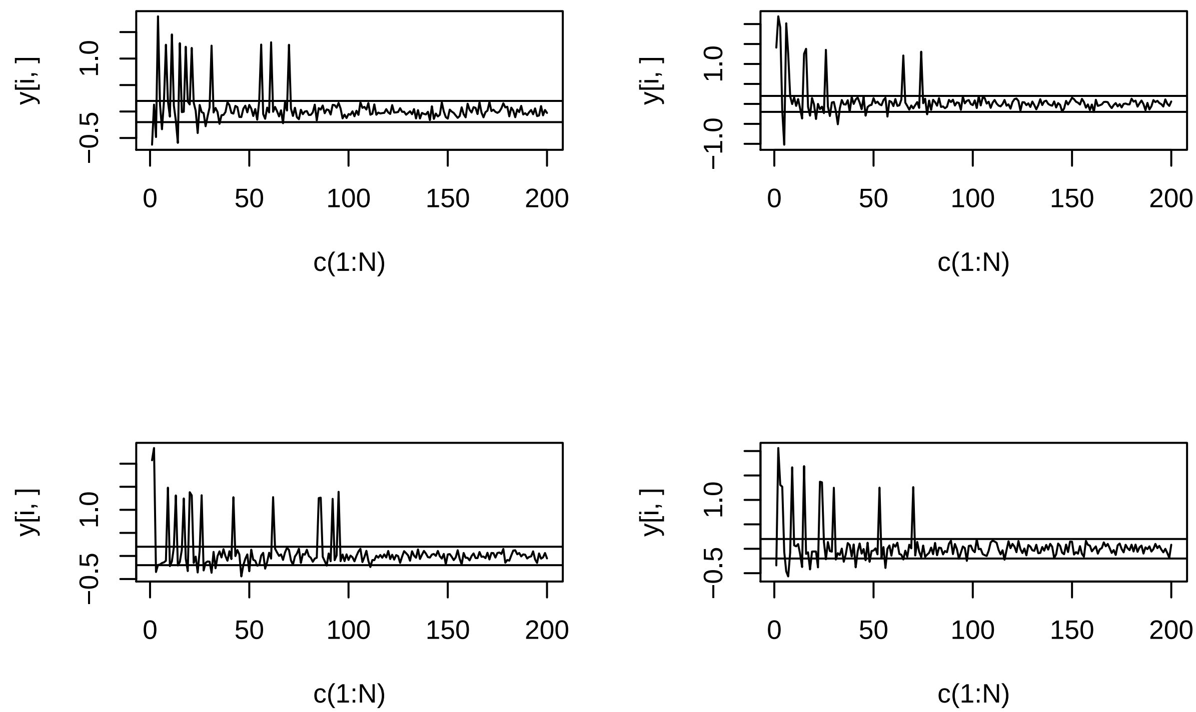
$$\frac{X_N}{\sqrt{N}}+1\sim N(1,\frac{1}{\sqrt{N}}) \tag{39}$$

Temos:

$$P[|Y_N|>\epsilon]=(1-[Q(\frac{-\epsilon-1}{\sqrt{\frac{1}{N}}})-Q(\frac{\epsilon-1}{\sqrt{\frac{1}{N}}})])Q(\frac{0.1}{\sqrt{\frac{1}{N}}})+(1-[Q(\frac{-\epsilon}{\sqrt{\frac{1}{N}}})-Q(\frac{\epsilon}{\sqrt{\frac{1}{N}}})])(1-Q(\frac{0.1}{\sqrt{\frac{1}{N}}})) \tag{40}$$

Como $\frac{1}{\sqrt{\frac{1}{N}}}=\sqrt{N}$, quando $N\to\infty$, ambos os termos tendem a 0 e, desta forma, $P[|Y_N|>\epsilon]\to 0$, C.Q.D..

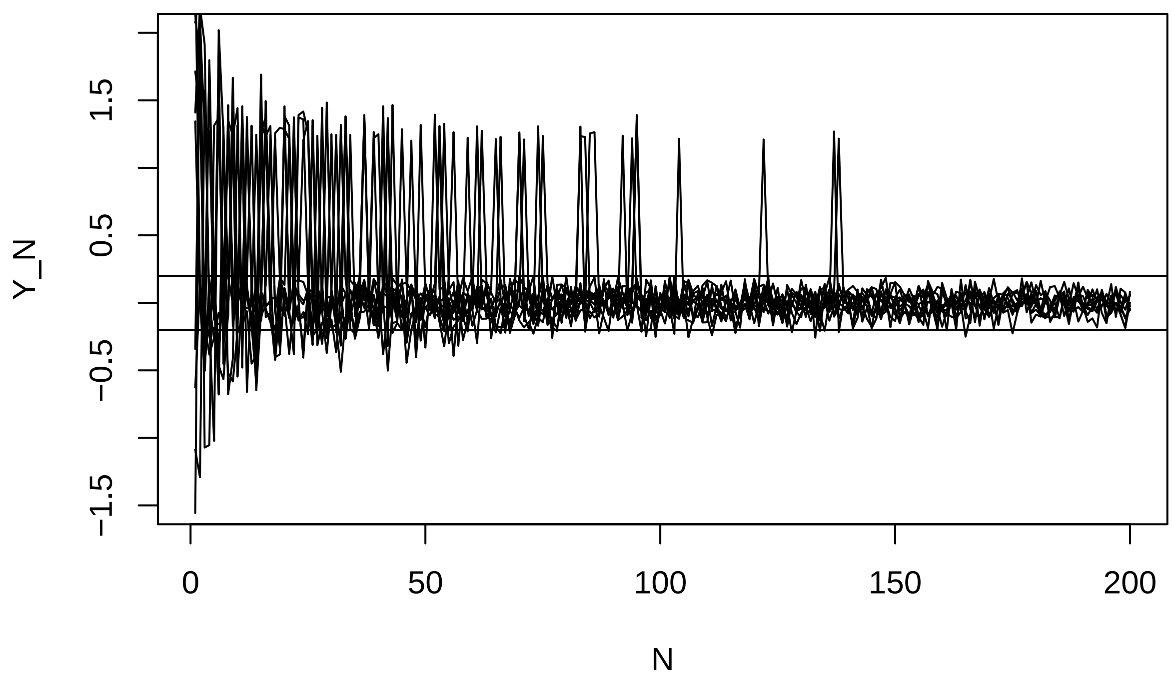
```
set.seed(1)

N<-200
R<-10

y<-matrix(0,R,N)

for (j in 1:R) {
        for (i in 1:N) {
                x<-rnorm(n=1)/sqrt(i)
                if (x<0.2) {
                        y[j,i]<-x
                } else {
                        y[j,i]<-x+1
                }
        }
}

par(mfrow=c(2,2))
for (i in 1:4) {
        plot(c(1:N),type="l",y[i,])
        abline(h=0.2);abline(h=-0.2)
}
```



Como podemos ver, para as 4 realizações, todas convergiram para N>100. No entanto, essa conclusão não pode ser generalizada para qualquer realização, conforme resultado da simulação, envolvendo 10 realizações.

O que é possível concluir é que a probabilidade de convergência é alta, para valores grandes de N.

```r
plot(c(1:N),type="l",y[1,],xlab="N",ylab="Y_N",xlim=c(1, N),ylim=c(-1.5,2))
abline(h=0.2);abline(h=-0.2)

for (j in 2:R) {
        par(new=T)
        plot(y[j,],type='l',xlab="",ylab="",axes=F,xlim=c(1, N),ylim=c(-1.5,2))
}
```



Exercício 15.9

Seja $S_N = \sum_{i=1}^{100} X_i$, em que X_i é a VA resistência. De acordo com o Teorema do Limite Central, S_N tem distribuição aproximadamente normal, com $E[S_N] = NE[X] = 1000$ e $var(S_N) = Nvar(X_i) = 200$.

Desta forma, a probabilidade da resitência dos 100 resistores conectados em série ser superior a 1030 é dada por:

$$P[S_N > 1030] \approx Q(\frac{1030 - 1000}{\sqrt(200)}) = 0.0169 \tag{41}$$

Exercício 15.12

A PDF de Y será aproximadamente normal, com média $E[Y]$ e variância $var(Y)$.

$$E[Y] = NE[X^2] = 100(var(X) + E^2[X]) = 100.24 = 2400 \tag{42}$$
$$\tag{43}$$
$$var(Y) = 100.var(X^2) = 100(E[X^2] - E^2[X^2]) \tag{44}$$
$$\tag{45}$$
$$E[X^2] = 24 \tag{46}$$

Para encontrar $E[X^4]$, define-se $Z \sim N(0, 8)$, tal que $X = Z - 4$. Desta forma,$E[X^4] = E[(Z-4)^4]$.

Desenvolvendo $E[(Z-4)^4]$ e sabendo que os momentos de ordem ímpar da pdf gaussiana com média 0 é igual a 0 e que o valor esperado de uma constante é igual a própria constante, temos que:

$$E[X^4] = E[(Z-4)^4 =] = E[Z^4] + 6E[V^2](-4)^2 + (-4)^2 = 3var(Z)^2 + 96var(V) + 256 = 1216 \tag{47}$$

Logo, $var(X^2) = 100(1216 - 576) = 640$, $var(Y) = 100var(x^2) = 6400$ e $Y \sim N(2400, 6400)$

Exercício 15.14

A PDF de Y será aproximadamente normal, com média $E[Y] = 10E[X]$ e variância $var(Y) = 10var(X)$.

$$E[X] = \int_{\infty}^{\infty} xp(x)dx = \int_0^1 2x^2dx = \frac{2}{3} \tag{48}$$

$$E[X^2] = \int_{\infty}^{\infty} x^2p(x)dx = \int_0^1 2x^3dx = \frac{1}{2} \tag{49}$$

$$var(X) = E[X^2] - E^2[X] = \frac{1}{2} - (\frac{2}{3})^2 = \frac{1}{18} \tag{50}$$

Logo, $Y \sim N(\frac{20}{3}, \frac{10}{18})$ e:

$$P(Y > 7) = Q(\frac{7 - \frac{20}{3}}{\sqrt{(\frac{5}{9})}}) = 0.3274 \tag{51}$$

Exercício 15.17

A distribuição chi-quadrado é um caso especial da pdf gamma $\Gamma(\alpha, \beta)$, com $\alpha = \frac{v}{2}$ e $\beta = 2\sigma^2$, em que v é a quantidade de graus de liberdade (igual a N nesse caso) e σ o desvio padrão da distribuição normal que derivou a chi-quadrado que, no caso da distribuição normalizada, equivale a 1.

Ou seja, $Y_N \sim \Gamma(\frac{N}{2}, 2)$ e sua função característica é dada por:

$$\Phi_{Y_N}(w) = \frac{1}{(1 - j\omega\beta)^{\alpha}} \tag{52}$$

Substituindo $\alpha = \frac{N}{2}$ e $\beta = 2$, temos que:

$$\Phi_{Y_N}(w) = \frac{1}{(1 - 2j\omega)^{\frac{N}{2}}} \tag{53}$$

Definindo $Z_N = \frac{Y_N - E[Y_N]}{\sqrt{var(Y_N)}}$ e sabendo que $E[Y_N] = N$ e $var(Y_N) = 2N$, temos que

$$Z_N = \frac{Y_N - N}{\sqrt{2N}} \tag{54}$$

A função característica de Z_N é dada por:

$$\Phi_{Z_N}(w) = E[e^{j\omega Z_N}] = E_{Y_N}[e^{j\omega \frac{Y_N - N}{\sqrt{2N}}}] \tag{55}$$
$$\tag{56}$$

Mas, separando $E_{Y_N}[e^{j\omega \frac{Y_N - N}{\sqrt{2N}}}]$ em seus dois termos, temos que:

$$E_{Y_N}[e^{j\frac{\omega}{\sqrt{(2N)}} Y_N}] = \Phi_{Y_N}(\frac{\omega}{\sqrt{(2N)}}) = \frac{1}{(1 - 2j\frac{\omega}{\sqrt{2N}})^{\frac{N}{2}}} \tag{57}$$
$$\tag{58}$$
$$E_{Y_N}[e^{-j\omega \frac{N}{\sqrt{2N}}}] = e^{\frac{-j\omega N}{\sqrt{2N}}} = e^{-j\omega \sqrt{\frac{N^2}{2N}}} = e^{-j\omega \sqrt{\frac{N}{2}}} \tag{59}$$

Logo:

$$\Phi_{Z_N} = \frac{e^{-j\omega \sqrt{\frac{N}{2}}}}{(1 - 2j\frac{\omega}{\sqrt{2N}})^{\frac{N}{2}}} \tag{60}$$

Tirando o logaritmo de ambos os lados da equação, temos que:

$$\ln \Phi_{Z_N} = -j\omega \sqrt{\frac{N}{2}} - \frac{N}{2} \ln{(1 - 2j\frac{\omega}{\sqrt{2N}})} \tag{61}$$

Com $N \to \infty$ e aplicando a aproximação dada no exercício, temos:

$$\ln \Phi_{Z_N} = -j\omega \sqrt{\frac{N}{2}} - \frac{N}{2}(-2j\frac{\omega}{\sqrt{2N}} - \frac{1}{2}(2j\frac{\omega}{\sqrt{2N}})^2) \tag{62}$$
$$\tag{63}$$
$$\ln \Phi_{Z_N} = -j\omega \sqrt{\frac{N}{2}} + j\omega \sqrt{\frac{N}{2}} + \frac{N}{4}(-\frac{2}{N}\omega^2) = -\frac{\omega^2}{2} \tag{64}$$

Logo, $\Phi_{Z_N} = e^{-\frac{\omega^2}{2}}$ e, consequentemnete, $Z_N \sim N(0, 1)$, C.Q.D..

Exercício 15.24

```
N<-1000
k1<-490
k2<-510

Y<-0

for (k in k1:k2) {
```

```
        lnpy=-N*log(2)

        for (i in (N-k+1):N) {
                lnpy=lnpy+log(i)
        }

        for (i in 1:k) {
                lnpy=lnpy-log(i)
        }

        Y<-Y+exp(lnpy)
}

print(Y)
```

```
## [1] 0.49334
```

O resultado encontrado é o mesmo dado pelo livro. Se o cálculo fosse feito diretamente para cada termo, o custo computacional e a complexidade do código seria bem maior.