Wiktoria Kaczmarek, Aktywność 03.

1. Zużycie wody (w hektolitrach) w pewnym osiedlu w ciągu dnia ma rozkład

$$\mathcal{N}(m = ?, \sigma = 11)$$

Obliczyć prawd. zdarzenia, że empiryczna wariancja zużycia wody w losowo wybranych 90 dniach

a) nie będzie większa niż 100[hl],
b) będzie większa niż 200[hl].

Odpowiedzi:

X_i – zmienna losowa opisująca zużycie wody w jednym losowo wybranym dniu, i=1,…,90

$\bar{X}$-rozkład średniej zużycia wody w 90 dniach

$$\bar{X} \sim N(m, \frac{11}{\sqrt{90}})$$

Oznaczmy Y, jako $X_i - \bar{X}$ wtedy Y ma rozkład normalny:

$$Y \sim N(m + m, \sqrt{\frac{\sigma^2}{n} + \sigma^2}) \text{ , } Y \sim N(m + m, \sqrt{\frac{121}{90} + 121})$$

Korzystając z twierdzenia: Jeżeli X_i, i=1,2,…,n są zmiennymi losowymi z rozkładem normalnym z parametrami μ=0, σ=1 to ich suma ma rozkład χ^2 z n stopniami swobody.

$$X_1 + X_2 + \cdots + X_n = \chi^2_n$$

Wariancja będzie miała następujący wzór:

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{90}(Y_i)^2$$

$$S_n^2 = \frac{1}{89}\sum_{i=1}^{90}(Y_i)^2 = \frac{1}{89}\sum_{i=1}^{90}\left(\frac{Y_i - \sqrt{90}}{\sqrt{91} \cdot 11}\right)^2 \cdot \frac{91 \cdot 11^2}{90} = \frac{91 \cdot 121}{90 \cdot 89}\chi^2_{90}$$

a) Obliczamy prawdopodobieństwo, że wariancja będzie mniejsza niż 100hl:

$$P(S_n^2 < 100) = P\left(\frac{91 \cdot 121}{90 \cdot 89}\chi^2_{90} < 100\right) = P\left(\chi^2_{90} < \frac{9000 \cdot 89}{91 \cdot 121}\right)$$

```
> pchisq(9000*121/(91*121),90)
[1] 0.7555559
```

WYNIK: 75,56%

b)

$$P(S_n^2 > 200) = 1 - P\left(\frac{91 \cdot 121}{90 \cdot 89}\chi^2_{90} < 200\right) = 1 - P\left(\chi^2_{90} < \frac{18000 \cdot 89}{91 \cdot 121}\right)$$

```
> 1-pchisq(18000*121/(91*121),90)
[1] 4.589235e-10
```

WYNIK=~0%

2. Wiadomo, że błąd pomiaru pewnego przyrządu ma rozkład normalny $\mathcal{N}(0,\sigma)$ i z prawd. 0,95 nie wychodzi poza przedział $(-1,1)$. Dokonanych zostanie i) 10, ii) 100 niezależnych pomiarów tym przyrządem. Oblicz prawd. zdarzenia, że wariancja pomiarów

a) przyjmie wartość między 0,2 a 0,3,

b) będzie większa od 0,28. Odp. i) a) 0,1665, ii) b) ≈ 0,27.

Odpowiedzi:

X_i – zmienna losowa opisująca błąd jednego pomiaru $X_i \sim N(0,\sigma)$

$$P(-1 < X_i < 1) = P\left(-\frac{1}{\sigma} < \frac{X_i}{\sigma} < \frac{1}{\sigma}\right) = \Phi\left(\frac{1}{\sigma}\right) - \Phi\left(-\frac{1}{\sigma}\right) = 2\cdot\Phi\left(\frac{1}{\sigma}\right) - 1 = 0.95$$

$$2\cdot\Phi\left(\frac{1}{\sigma}\right) - 1 = 0.95$$

$$\Phi\left(\frac{1}{\sigma}\right) = 0.975$$

$$\left(\frac{1}{\sigma}\right) = \Phi^{-1}(0.975)$$

$$\sigma = \frac{1}{\Phi^{-1}(0.975)}$$

```
> 1/qnorm(0.975,0,1)
[1] 0.5102135
```

a) Wariancja próby wyznacza się za pomocą wzoru:

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X_n})^2$$

Rozkład średniej: $X_n \sim N(\mu, \frac{\sigma}{\sqrt{n}})$

Oznaczmy Y, jako $X_i - \bar{X}$ wtedy Y ma rozkład normalny:

$$Y \sim N(0+0, \sqrt{\frac{\sigma^2}{n} + \sigma^2})$$

$$Y \sim N(0+0, \sqrt{\frac{n+1}{n}\sigma})$$

Korzystając z twierdzenia: Jeżeli X_i, i=1,2,…,n są zmiennymi losowymi z rozkładem normalnym z parametrami μ=0, σ=1 to ich suma ma rozkład χ^2 z n stopniami swobody.

$$X_1 + X_2 + \cdots + X_n = \chi_n^2$$

Wariancja będzie miała następujący wzór:

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i)^2$$

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{Y_i \cdot \sqrt{n}}{\sqrt{n+ln\sigma}}\right)^2 \cdot \frac{(n+1)\sigma^2}{n} = \frac{(n+1)\sigma^2}{n(n-1)}\chi_n^2$$

$$P(0,2 < S_n^2 < 0,3) = ?$$

$$P(0,2 < S_n^2 < 0,3) = P\left(0,2 < \frac{(n+1)\sigma^2}{n(n-1)}\chi_n^2 < 0,3\right) = P\left(0,2\frac{n(n-1)}{(n+1)\sigma^2} < \chi_n^2 < 0,3\frac{n(n-1)}{(n+1)\sigma^2}\right)$$

i) n=10

$$P\left(0,2\frac{10(10-1)}{(10+1)\sigma^2} < \chi_n^2 < 0,3\frac{10(10-1)}{(10+1)\sigma^2}\right)$$

```
> odch<- 1/qnorm(0.975,0,1)
> pchisq(0.3*90/(11*odch^2),10)-pchisq(0.2*90/(11*odch^2),10)
[1] 0.2987606
```

ii)

$$P\left(0,2\frac{100(100-1)}{(100+1)\sigma^2} < \chi_n^2 < 0,3\frac{100(100-1)}{(100+1)\sigma^2}\right)$$

```
> odch<- 1/qnorm(0.975,0,1)
> pchisq(0.3*9900/(101*odch^2),100)-pchisq(0.2*9900/(101*odch^2),100)
[1] 0.7918891
```

b)

$$P(S_n^2 > 0,28) = 1 - P\left(\frac{(n+1)\sigma^2}{n(n-1)}\chi_n^2 < 0,28\right) = 1 - P\left(\chi_n^2 < 0,28\frac{n(n-1)}{(n+1)\sigma^2}\right)$$

i) n=10

$$P(S_n^2 > 0,28) = 1 - P\left(\chi_n^2 < 0,28\frac{10(10-1)}{(10+1)\sigma^2}\right)$$

```
> odch<- 1/qnorm(0.975,0,1)
> 1-pchisq(0.28*90/(11*odch^2),10)
[1] 0.5511423
```

ii) n=100

```
> odch<- 1/qnorm(0.975,0,1)
> 1-pchisq(0.28*9900/(101*odch^2),100)
[1] 0.335696
```

UWAGA: W wykładzie znajdował się inny wzór, lecz analizując wiele innych źródeł wybrałam wzór zależności, który według mojej intuicji był najbardziej dokładny, stąd moje wyniki odbiegają od Pana podanych w odpowiedzi.

3. Losujemy 100 liczb według rozkładu jednostajnego na przedziale $(0,1)$.

a) Ustalić rozkład sumy tych liczb.

b) Obliczyć prawd. zdarzenia, że suma wylosowanych liczb nie będzie należała do przedziału $(45,55)$.

c) Wyznaczyć dystrybuantę największej z wylosowanych liczb i oblicz prawd., że liczba ta będzie mniejsza od 0,95.

d) Jaki wniosek należy wyciągnąć, jeśli że suma wylosowanych liczb będzie mniejsza niż 40?

e) Przeprowadzić eksperyment losowania i porównać z wynikami teoretycznymi.

Odpowiedzi:

a) $X_i \sim U(0,1)$

Funkcja gęstości: $f_{X_i}(x) = \begin{cases} 0, & x \le 0 \lor x \ge 1 \\ 1, & 0 < x < 1 \end{cases}$

Dystrybuanta: $F_{X_i}(t) = \begin{cases} 0, & t \le 0 \\ t, & 0 < t < 1 \\ 1, & t \ge 1 \end{cases}$

Korzystając z CTG dla sumy, musimy założyć, że n=100 jest wystarczająca duża liczbą, wtedy:

$$X = \sum_{i=1}^{100} X_i \sim N(n \cdot EX, \sqrt{n} \cdot DX)$$

Gdzie EX jest wartością oczekiwaną, czyli $\frac{b-a}{2} = \frac{1}{2}$,

a DX jest odchyleniem standardowym $DX = \sqrt{\frac{(b-a)^2}{12}} = \frac{1}{\sqrt{12}}$.

Zatem $X \sim N(50, \frac{5}{\sqrt{3}})$

b) $P(x < 45 \lor X > 55) = ?$

$$P(x < 45 \lor X > 55) = 1 - P(45 < X < 55) = 1 - (F(55) - F(45))$$

```
> odch<- 5/sqrt(3)
> 1-(pnorm(55,50,odch)-pnorm(45,50,odch))
[1] 0.08326452
```

c) $Y = \sum_{i=1} 100 X_i$

$$F_Y(t) = P(Y \le t) = \prod_{i=1}^{100} P(X_i \le t)$$

$$\prod_{i=1}^{100} P(X_i \le t) = P(X_i \le t)^{100} = \begin{cases} 0, & t \le 0 \\ t^{100}, & 0 < t < 1 \\ 1, & t \ge 1 \end{cases}$$

$$P(Y < 0.95) = F_Y(0.95) = 0.95^{100}$$

```
> 0.95^100
[1] 0.005920529
```

d) P(X<40)=?

```
> odch<- 5/sqrt(3)
> pnorm(40,50,odch)
[1] 0.0002660028
```

Jeżeli suma z wylosowanych liczb będzie mniejsza od 40 to oznacza, że wartość oczekiwana jest zawyżona lub przybliżenie jest błędne. Prawdopodobieństwo wystąpienie tej sytuacji jest małe.

5. Ze zbioru $\{1, 2, 3, 4\}$ wylosowano dwie liczby

i) ze zwracaniem,

ii) bez zwracania,

Wyznaczyć rozkład oraz wartość oczekiwaną i wariancję rozstępu.

Odpowiedz:

a) Niech X_1, X_2 będą losowane ze zbioru {1,2,3,4}.

$\Omega = 4^2 = 16$

A=(X_1, X_2)

R – rozstęp

Tabela 1. Tabela opisująca Ω.

A	$P(X_1, X_2)$	$X_{(1)}$	$X_{(2)}$	R
(1,1)	1/16	1	1	0
(1,2)	1/16	1	2	1
(1,3)	1/16	1	3	2
(1,4)	1/16	1	4	3
(2,1)	1/16	1	2	1
(2,2)	1/16	2	2	0
(2,3)	1/16	2	3	1
(2,4)	1/16	2	4	2
(3,1)	1/16	1	3	2
(3,2)	1/16	2	3	1
(3,3)	1/16	3	3	0
(3,4)	1/16	3	4	1
(4,1)	1/16	1	4	3
(4,2)	1/16	2	4	2
(4,3)	1/16	3	4	1
(4,4)	1/16	4	4	0

Tabela 2. Tabela przedstawiająca rozkład łączny i rozkłady brzegowe.

$x_1 \backslash x_2$	1	2	3	4	$f_{X(1)}(x_1)$
1	1/16	2/16	2/16	2/16	7/16
2	0	1/16	2/16	2/16	5/16
3	0	0	1/16	2/16	3/16
4	0	0	0	1/16	1/16
$f_{X(2)}(x_2)$	1/16	3/16	5/16	7/16	16/16

$$E(X_{(1)}) = \sum_{i=1}^{4} x_i \cdot f_{X(1)}(x_i) = \frac{7}{16} + \frac{10}{16} + \frac{9}{16} + \frac{4}{16} = \frac{15}{8}$$

$$E(X_{(1)}^2) = \sum_{i=1}^{4} x_i^2 \cdot f_{X(1)}(x_i) = \frac{7}{16} + \frac{20}{16} + \frac{27}{16} + \frac{16}{16} = \frac{35}{8}$$

$$D^2(X_{(1)}) = E(X_{(1)}^2) - E(X_{(1)})^2 = \frac{35}{8} - \frac{225}{64} = \frac{55}{64}$$

$$E(X_{(2)}) = \sum_{i=1}^{4} x_i \cdot f_{X(2)}(x_i) = \frac{1}{16} + \frac{6}{16} + \frac{15}{16} + \frac{28}{16} = \frac{25}{8}$$

$$E(X_{(2)}^2) = \sum_{i=1}^{4} x_i^2 \cdot f_{X(2)}(x_i) = \frac{1}{16} + \frac{12}{16} + \frac{35}{16} + \frac{112}{16} = \frac{85}{8}$$

$$D^2(X_{(2)}) = E(X_{(2)}^2) - E(X_{(2)})^2 = \frac{85}{8} - \frac{625}{64} = \frac{55}{64}$$

$$E(X_{(1)}X_{(2)}) = \sum_{i=1}^{4}\sum_{j=1}^{4} x_i \cdot x_j \cdot f_X(x_i, x_j) = \frac{19}{16} + \frac{32}{16} + \frac{33}{16} + \frac{16}{16} = \frac{25}{4}$$

$$Cov(X_{(1)}X_{(2)}) = E(X_{(1)}X_{(2)}) - E(X_{(1)}) \cdot E(X_{(2)}) = \frac{400 - 375}{64} = \frac{35}{64}$$

Tabela 3. Tabela przedstawiająca prawdopodobieństwa dla wartości rozstępów.

R_i	0	1	2	3
$P(R_i)$	4/16	6/16	4/16	2/16

$$ER = \sum_{i=0}^{3} R_i \cdot P(R_i) = 0 + \frac{6}{16} + \frac{8}{16} + \frac{6}{16} = \frac{5}{4}$$

$$E(R^2) = \sum_{i=0}^{3} R_i^2 \cdot P(R_i) = 0 + \frac{6}{16} + \frac{16}{16} + \frac{18}{16} = \frac{5}{2}$$

$$D^2R = E(R^2) - [ER]^2 = \frac{15}{16}$$

b) Analogicznie do a) tworzymy dwie tabele.
Niech X_1, X_2 będą losowane ze zbioru {1,2,3,4}.
$\Omega = 12$
A=(X_1, X_2)
R – rozstęp

Tabela 4. Tabela opisująca Ω.

A	P(X_1, X_2)	$X_{(1)}$	$X_{(2)}$	R
(1,2)	1/12	1	2	1
(1,3)	1/12	1	3	2
(1,4)	1/12	1	4	3
(2,1)	1/12	1	2	1
(2,3)	1/12	2	3	1
(2,4)	1/12	2	4	2
(3,1)	1/12	1	3	2
(3,2)	1/12	2	3	1
(3,4)	1/12	3	4	1
(4,1)	1/12	1	4	3
(4,2)	1/12	2	4	2
(4,3)	1/12	3	4	1

Tabela 5. Tabela przedstawiająca rozkład łączny i rozkłady brzegowe.

$x_1 \backslash x_2$	2	3	4	$f_{x(1)}(x_1)$
1	2/12	2/12	2/12	2/12
2	0	2/12	2/12	2/12
3	0	0	2/12	2/12
$f_{x(2)}(x_2)$	2/12	4/12	6/12	12/12

Tabela 6. Tabela przedstawiająca prawdopodobieństwa dla wartości rozstępów.

R_i	1	2	3
$P(R_i)$	6/12	4/12	2/12

$$ER = \sum_{i=0}^{3} R_i \cdot P(R_i) = \frac{6}{12} + \frac{8}{12} + \frac{6}{12} = \frac{5}{3}$$

$$E(R^2) = \sum_{i=0}^{3} R_i^2 \cdot P(R_i) = \frac{6}{12} + \frac{14}{12} + \frac{18}{12} = \frac{10}{3}$$

$$D^2R = E(R^2) - [ER]^2 = \frac{5}{9}$$