1. Dwuwymiarowy rozkład pary zmiennych losowych X oraz Y dany jest za pomocą tablicy

X \ Y	-1	1
-5	0.2	0.1
1	0.3	0.4

Wyznaczyć $E(-4X - 2Y^2)$ oraz $E(XY)$.

$$P(x=0 \mid y<2) = \frac{P(x=0; y<2)}{P(y<2)}$$

2. Pewna partia polityczna ma dwudziestoprocentowe poparcie w społeczeństwie. Jakie jest prawdopodobieństwo, że w okręgu wyborczym składającym się z 10000 wyborców, partia uzyska co najmniej 1000 głosów przy siedemdziesięcioprocentowej frekwencji. Zastosuj przybliżenie rozkładem normalnym lub Poissona.

3. Sklep spożywczy zaopatrują w spirytus dostawcy D_1, D_2, D_3 zachowując proporcje $1:1:2$. Wiadomo, że dostawy D_1 zawierają średnio 2% butelek z uszkodzonym stemplem lakowym, dostawy D_2 – 4%, a dostawy D_3 – 1%. Klient, który kupił butelkę z uszkodzonym stemplem lakowym kwestionuje jej zawartość. Wyznaczyć prawdopodobieństwo, że:

a) na tysiąc reklamowanych butelek co najmniej 300 pochodzi od dostawcy D_3,

b) na sto butelek, co najmniej cztery będą reklamowane.

Ponadto zbadać, ile co najmniej butelek należy kupić, aby z prawdopodobieństwem nie mniejszym niż 0.99 trafić na co najmniej jedną butelkę z uszkodzonym stemplem lakowym.

W punktach a) i b) zastosować przybliżenie rozkładem Poissona lub normalnym uzasadniając swój wybór.

4. Wśród tysiąca skontrolowanych w centralnej Polsce pojazdów osiemdziesiąt nie posiadało ubezpieczenia. W podobnej kontroli przeprowadzonej w zachodniej części kraju na półtora tysiąca pojazdów sto czterdzieści nie posiadało ubezpieczenia. Ocenić czy odsetek pojazdów bez ubezpieczenia zależy od części kraju.

5. Badano dzienny czas poświęcany przez dzieci w wieku przedszkolnym na oglądaniu telewizji. Uzyskano następujące wyniki na próbie przedszkolaków (w minutach): 132, 114, 51, 97, 117, 119, 122, 65, 109, 84, 85, 134, 133, 107, 149. Czy zebrane wyniki potwierdzają przypuszczenia, że przedszkolaki spędzają na oglądaniu telewizji przeciętnie dwie godziny dziennie?
Obliczenia: $\bar{x} = 107.8667$, $s = 27.14739$

6. Postanowiono oszacować parametr CCS (Czytelnictwo Cyklu Sezonowego) dla pewnego kwartalnika. Ponieważ jest to kwartalnik, za sezon przyjmuje się jeden rok. Zatem interesuje nas odsetek osób, które miało kontakt z danym kwartalnikiem w ciągu roku. Okazało się, że na tysiąc wylosowanych osób z grupy docelowej, 350 miało kontakt z tym kwartalnikiem. Oszacować interesujący nas parametr CCS na poziomie ufności 95%. Zinterpretować wynik.

7. Z dotychczasowych obserwacji wynika, że zainteresowanie mężczyzn pewnym kierunkiem studiów jest niewielki i można je szacować na poziomie 1%. Jakie jest prawdopodobieństwo, że wśród 200 osób przyjmowanych w kolejnej rekrutacji będzie co najmniej trzech mężczyzn? Zastosuj przybliżenie rozkładem normalnym lub Poissona.

8. Dwuwymiarowy rozkład pary zmiennych losowych X oraz Y dany jest za pomocą tablicy

X \ Y	-1	1
-3	0.1	0.2
1	0.3	0.4

Wyznaczyć $E(X - 2Y^2)$ oraz $E(Y - XY)$.

9. Sklep zaopatrują w towar trzy zakłady odpowiednio w dwudziestu, pięćdziesięciu oraz trzydziestu procentach. W dostawach tych zakładów procent braków jest odpowiednio równy: 0.5%, 0.3% oraz 0.5%. Wyznaczyć prawdopodobieństwo, że:

a) na tysiąc zakupionych sztuk towaru co najmniej 300 pochodzi z drugiego zakładu,

b) na sto zakupionych sztuk towaru co najwyżej cztery, to braki.

Ponadto zbadać, ile co najmniej sztuk towaru trzeba zakupić, aby z prawdopodobieństwem nie mniejszym niż 0.7 trafić na co najmniej jedną wybrakowaną sztukę towaru z zakładu drugiego.

W punktach a) i b) zastosować przybliżenie rozkładem Poissona lub normalnym uzasadniając swój wybór.

10. Z pewnego regionu Polski wylosowano do badań tysiąc osób, w tym dwieście osób z wyższym wykształceniem. Wśród osób z wyższym wykształceniem zliczono dziesięć osób bezrobotnych, a wśród osób bez wyższego wykształcenia sześćdziesiąt cztery. Czy na podstawie uzyskanych rezultatów można uznać, że odsetek bezrobotnych zależy od poziomu wykształcenia?

11. Badano czy na zmianę wydajności pracowników wpływa przebyte przeszkolenie. Eksperyment przeprowadzono na wybranych losowo pracownikach. Zmiany ich wydajności po przeszkoleniu były następujące: -12, 13, -11, 14, 15, 9, -5, 5, 4, 8, 14, 6. Czy uzyskane wyniki potwierdzają przypuszczenie, że badany rodzaj szkolenia nie ma istotnego wpływu na zmianę wydajności?
Obliczenia: $\bar{x} = 5$, $s = 9.53$

12. Postanowiono oszacować parametr *OTS* 1+ (Opportunity To See) określający zasięg kampanii reklamowej: odsetek osób, które przynajmniej raz miały styczność z daną reklamą w czasie trwania kampanii reklamowej. Z grupy docelowej wybrano losowo 1000 osób; 850 spośród nich miały styczność z reklamą. Na poziomie ufności 95% oszacować *OTS* 1+. Zinterpretować wynik.

(1.)

x \ Y	-1	1
-5	0,2	0,1
1	0,3	0,4

Wyznaczyć $E(-4X-2Y^2)$ oraz $E(XY)$

$\mathbb{P}(x=-5) = 0,2 + 0,1 = 0,3$

$\mathbb{P}(x=1) = 0,3 + 0,4 = 0,7$

$\mathbb{P}(y=-1) = 0,2 + 0,3 = 0,5$

$\mathbb{P}(y=1) = 0,1 + 0,4 = 0,5$

$EX = \cancel{-5 \cdot 0,3 = -1,5}\ -5 \cdot 0,3 + 0,7 = -0,8$

$Ey = -1 \cdot 0,5 + 0,5 = 0$ dla $y=1$ $Ey = 0,5$

$E(XY) = (-5 \cdot -1 \cdot 0,2) + (-5 \cdot 0,1) + (-1 \cdot 0,3) + 0,4$

$[E(XY) = 1 + (-0,5) + (-0,3) + 0,4 = 0,6]$

$E(-4x-2y^2) = -4EX - 2Ey^2$

$E(-4x-2y^2) = -4 \cdot -0,8 - 2*1=3,2 - 2=1,2$

EY^2=(-1)^2 *0,5 + 1^2 *0,5=1

(2.)

$n = 0,7 \cdot 10000 = 7000$

$p = 0,2$ $q = 1-p = 0,8$

$P(X \geq 1000) = ?$

Zmienną losową X przybliżamy rozkładem normalnym ze średnią $np = 1400$ i wariancją $npq = 1120$

$\mathbb{P}(X \geq 1000) = \mathbb{P}\left(\frac{X-1400}{\sqrt{1120}} \geq \frac{1000-1400}{\sqrt{1120}}\right)$

Zmienna losowa $\frac{X-1400}{\sqrt{1120}}$ ma rozkład st. normalny i jego dystrybuantę odczytujemy z tablic.

Zatem $\mathbb{P}\left(\underbrace{\frac{X-1400}{\sqrt{1120}}}_{= Z} \leq 11,95\right) = \mathbb{P}(Z \leq 11,95) = 1 - \Phi(11,95)$

$= 1 - 0,88298 = 0,11702$

③ D_1 – towar od I ~~zakładu~~ dostawcy
D_2 – // – II
D_3 – // – III

0,25 0,25 0,5
$1:1:2$

$P(D_1) = 0{,}25$
$P(D_2) = 0{,}25$
$P(D_3) = 0{,}5$
} szanse w jednym losowaniu

B – wylosowanie ~~butelki~~ butelki z uszkodzonym stemplem lakowym

$P(B|D_1) = 0{,}02$ – wylosowanie uszkodzonej perełki lakowej od I ~~zakładu~~ dostawcy
$P(B|D_2) = 0{,}04$ – // – // II
$P(B|D_3) = 0{,}01$ – // – III

a) X – ilość wylosowanych sztuk od III dostawcy
$p = 0{,}5$ – sukces = wylosowanie butelki od III dostawcy
$n = 1000$
$\mathbb{P}(X \geq 300) = ?$

Zastosujemy przybliżenie rozkładem normalnym (bo $np = 500$ jest duże)

$$\mathbb{P}(X \geq 300) = \mathbb{P}\left(\frac{X-500}{\sqrt{0{,}5 \cdot 1000 \cdot 0{,}5}} \geq \frac{300-500}{\sqrt{250}}\right)$$

npq = wariancja

$$Z = \frac{X-500}{\sqrt{250}} \sim N(0,1)$$

$= \mathbb{P}(Z \geq -1{,}265) = \mathbb{P}(Z \leq 1{,}265)$ – otrzymujemy z tablic

$$\mathbb{P}(Z \leq 1{,}265) = 1 - \Phi(1{,}265) = 1 - 0{,}89617 = 0{,}10383$$

b) $p = P(B) = P(B|D_1) \cdot P(D_1) + P(B|D_2) \cdot P(D_2) + P(B|D_3) \cdot P(D_3) =$
$= 0{,}02$ $\quad n = 100$
$np = 2$

↖ wzór nie przeor. całkowite.

X – ilość uszkodzonych – przybliżamy rozkładem Poissona z par $\lambda = 2$

$$\mathbb{P}(X \geq 4) = e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \frac{\lambda^4}{4!}\right) = \cancel{2{,}718^{-2}\left(1 + 0{,}4 + \frac{2^2}{2!} + \frac{2^3}{3!} + \frac{2^4}{4!}\right)}$$

$e = 2{,}718$

musimy stosować rozkład poissona bo np jest małe

$\cancel{\mathbb{P}(X \geq 4) = 0{,}999}$ $\quad \mathbb{P}(X \geq 4) = 1 - 0{,}9469 = 0{,}0531$

c) X – ilość butelek uszkodzonych od wszystkich ~~zakładów~~ dostawców
$p = P(B) = 0{,}02$
n – nieznane
$\mathbb{P}(X \geq 1) = 0{,}99 = 1 - P(X=0)$
$0{,}01 = P(X=0) = e^{-\lambda}$
$\lambda = -\ln(0{,}01) \approx 4{,}605$

$$\boxed{n = \frac{\lambda}{p} = 230}$$

(4). Dane

$n_1 = 1000$ $n_2 = 1500$

$m_1 = 80$ $m_2 = 140$

$\alpha = 0,05$

Test hipotezy o dwóch wskaźnikach struktury...

1) $H_0: p_1 = p_2$

$H_1: p_1 \neq p_2$

2) Statystyka badawcza

$$T = \frac{p_1 - p_2}{\sqrt{\frac{\bar{p} \cdot \bar{q}}{n}}} \qquad p_1 = \frac{m_1}{n_1} = 0,08 \qquad p_2 = \frac{m_2}{n_2} = 0,093$$

$$\bar{p} = \frac{m_1 + m_2}{n_1 + n_2} = \frac{80 + 140}{1000 + 1500} = 0,088 \qquad n = \frac{n_1 \cdot n_2}{n_1 + n_2} = \frac{1000 \cdot 1500}{2500} = 600$$

$$\bar{q} = 1 - \bar{p} = 1 - 0,088 = 0,912$$

Statystyka ta, przy założeniu prawdziwości hipotezy H_0 ma rozkład zbieżny do $N(0,1)$

$$T = \frac{0,08 - 0,093}{\sqrt{\frac{0,08 \cdot 0,912}{600}}} = -1,12$$

Obszar krytyczny $K = (-\infty; -1,96) \cup (1,96; \infty)$

$\Phi(\emptyset) = 1 - \frac{\alpha}{2} = 0,975$ $z_{0,975} = 1,96$

$T \notin K$ brak podstaw do odrzucenia H_0 o tym że odsetek pojazdów bez ubezpieczenia nie zależy od części kraju.

(5). Dane

$\bar{x} = 107{,}8667$ $n = 15$

$s = 27{,}14739$

$\alpha = 0{,}01$

1) $H_0: m = 120$

$H_1: m \neq 120$

2) Statystyka badawcza

Ponieważ $n < 30$ i σ nieznane zatem sprawdzianem jest $T_{n-1} = \frac{\bar{x} - m}{s} \cdot \sqrt{n-1}$, który ma rozkład Studenta o 14 stopniach swobody

$$T_{n-1} = \frac{107{,}8667 - 120}{27{,}14739} \cdot \sqrt{14} = -1{,}6723$$

Z rozkładu Studenta przy 14 st swobody i $\alpha = 0{,}05$ otrzymujemy $t_{\alpha, n-1} = 2{,}1448$

Obszar krytyczny $K = (-\infty; -2{,}1448) \cup (2{,}1448; \infty)$

$T_{n-1} \notin K$ obliczona statystyka wpada w obszar krytyczny.

Brak podstaw do odrzucenia hipotezy zerowej. Można uważać, że przeciętnie przedszkolaki spędzają przed TV 2 godziny

6. Przedział ufności dla wskaźnika struktury

$$P\left\{\frac{m}{n} - z_\alpha \sqrt{\frac{\frac{m}{n}(1 - \frac{m}{n})}{n}} < p_0 < \frac{m}{n} + z_\alpha \sqrt{\frac{(1)}{n}}\right\} = 1 - \alpha$$

Dane: $m = 350$

$n = 1000$

$1 - \alpha = 0{,}95$

$\Phi(z_\alpha) = 1 - \frac{\alpha}{2} = 0{,}975$

$z_\alpha = 1{,}96$

$$P\left\{0{,}35 - 1{,}96 \cdot \sqrt{\frac{0{,}35 \cdot 0{,}65}{1000}} < p_0 < 0{,}35 + 1{,}96 \sqrt{\frac{0{,}35 \cdot 0{,}65}{1000}}\right\} = 0{,}95$$

$$P\{0{,}3349 < p_0 < 0{,}3651\} = 0{,}95$$

Z prawdopodobieństwem 95% przedział ufności od 33,49% do 36,51% obejmuje odsetek osób, które miały kontakt z [illegible].

7)

$P(X \geq 3)$ $\quad n = \frac{\lambda}{p}$

$p = 0{,}01$ $\quad \lambda = n \cdot p = 2$

$n = 200$

$q = 0{,}99$

~~np=2 jest małe~~

x – ilość mężczyzn w kolejnej rekrutacji

– przybliżam rozkładem Poissona z par $\lambda = 2$

$$P\{X \geq 3\} = e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!}\right) = 2{,}718^{-2}\left(1 + 2 + 2 + \frac{8}{6}\right) = 0{,}1353 \cdot 6{,}33 = 0{,}856$$

85%

8)

X \ Y	-1	1
-3	0,1	0,2
1	0,3	0,4

Wyznacz $E(X - 2Y^2)$ oraz $E(Y - XY)$

$P(X = -3) = 0{,}1 + 0{,}2 = 0{,}3$

$P(X = 1) = 0{,}3 + 0{,}4 = 0{,}7$

$P(Y = -1) = 0{,}1 + 0{,}3 = 0{,}4$

$P(Y = 1) = 0{,}2 + 0{,}4 = 0{,}6$

$EX = -3 \cdot 0{,}3 + 0{,}7 = -0{,}2$

$EY = -0{,}4 + 0{,}6 = 0{,}2$

$EXY = (-3 \cdot -1 \cdot 0{,}1) + (-3 \cdot 0{,}2) + (-1 \cdot 0{,}3) + 0{,}4 = 0{,}3 + (-0{,}6) - 0{,}3 + 0{,}4$

$EXY = -0{,}2$

EY^2=0,4+0,6=1

$E(X - 2Y^2) = EX - 2EY^2 = -0{,}2 - 2 \cdot$ 1= -2,2

$E(Y - XY) = EY - EXY = 0{,}2 + 0{,}2 = 0{,}4$

9.

D_1 - towar od I ~~dostawy~~ zakładu $\quad P(D_1) = 0{,}2$
D_2 - -//- II $\quad P(D_2) = 0{,}5$ $\quad$ szanse wylosowania
D_3 -//- III $\quad P(D_3) = 0{,}3$

B - ~~braki~~ wylosowanie braku

$P(B|D_1) = 0{,}05$ - wylosowanie braku, jeżeli losujemy od I ~~dostawy~~ zakładu
$P(B|D_2) = 0{,}03$ -//- II
$P(B|D_3) = 0{,}05$ -//- III

a)
x - ilość wylosowanych sztuk od II zakładu

$P(x \geq 300) = ?$

$p = 0{,}5$ - sukces - wylosowanie ~~braku~~ towaru od II zakładu

$n = 1000 \quad q = 0{,}5$

Zastosujemy przybliżenie rozkładem normalnym (bo $np = 500$ jest duże)

$$P(x \geq 300) = P\left(\frac{x-500}{0{,}5 \cdot 1000 \cdot 0{,}5} \geq \frac{300-500}{\sqrt{250}}\right) \quad npq \text{ - wariancja}$$

$$Z = \frac{x-500}{\sqrt{250}} \sim N(0,1)$$

$$= P(Z \leq 1{,}265) = 1 - \Phi(1{,}265) = 1 - 0{,}89617 = 0{,}10383$$

b) $p = P(B) = P(B|D_1) \cdot P(D_1) + P(B|D_2) \cdot P(D_2) + P(B|D_3) \cdot P(D_3) = 0{,}01 + 0{,}015 + 0{,}015 = 0{,}04$

$n = 100 \quad np = 4$

x - ilość braków - przybliżamy rozkładem Poissona z par $\lambda = 4$ (np małe)

$$P(x \leq 4) = e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \frac{\lambda^4}{4!}\right) = 2{,}718^{-4} \cdot (1 + 4 + 8 + 5{,}333 + 5{,}333) = 0{,}01832 \cdot (23{,}666) = 0{,}433$$

c)
x - ilość braków z zakładu II

$p = P(B|D_2) = 0{,}03$

$P(x \geq 1) = 0{,}7 = 1 - P(x = 0)$

$0{,}3 = P(x = 0) = e^{-\lambda}$

$\lambda = -\ln(0{,}3) \approx 1{,}203 \qquad n = \frac{\lambda}{p} = \frac{1{,}203}{0{,}03} \approx 40$

10.

Dane

~~$n_1 = 1000$~~ ~~$n_2 = 64$~~

~~$m_1 = 200$~~ ~~$m_2 = 10$~~

$n_1 = 800$ $n_2 = 200$

$m_1 = 64$ $m_2 = 10$

~~$\alpha = 0{,}05$~~

Test hipotezy o dwóch wskaźnikach struktury

1) $H_0: p_1 = p_2$

$H_1: p_1 \neq p_2$

statystyka badawcza

$$T = \frac{p_1 - p_2}{\sqrt{\frac{\bar{p} \cdot \bar{q}}{n}}} \qquad p_1 = \frac{m_1}{n_1} = 0{,}08$$

$$p_2 = \frac{m_2}{n_2} = 0{,}05$$

$$\bar{p} = \frac{m_1 + m_2}{n_1 + n_2} = \frac{74}{1000} = 0{,}074 \qquad n = \frac{n_1 \cdot n_2}{n_1 + n_2} = \frac{800 \cdot 200}{1000} = 160$$

$$\bar{q} = 1 - \bar{p} = 0{,}926$$

statystyka ta, przy założeniu prawdziwości hipotezy H_0 ma rozkład zbliżony do $N(0,1)$

$$T = \frac{0{,}08 - 0{,}05}{\sqrt{\frac{0{,}074 \cdot 0{,}926}{160}}} = \frac{0{,}03}{0{,}02069} = 1{,}4499 \approx 1{,}45$$

Obszar krytyczny $K = (-\infty, -1{,}45) \cup (1{,}45, +\infty)$

$$z(\Phi) = 1 - \frac{\alpha}{2} = 0{,}926 \qquad z_{0{,}926} = 1{,}45$$

$$\alpha = 0{,}148$$

$T \notin K$ brak podstaw do odrzucenia H_0 o tym, że odsetek bezrobotnych nie zależy od poziomu wykształcenia

(11)

Dane:

$\bar{x} = 5$ $\quad n = 12$

$S = 9{,}53$

1) $H_0: m = 100$

$H_1: m = 120$

2) statystyka badawcza

Ponieważ $n < 30$ i σ nieznane zatem sprawdzianem jest $T_{n-1} = \frac{\bar{x} - m}{s}\sqrt{n-1}$, który ma rozkład studenta o 11 stopniach swobody

$T_{n-1} =$

---

12 Przedział ufności dla wskaźnika struktury

$$P\left\{\frac{m}{n} - Z_\alpha \sqrt{\frac{\frac{m}{n}(1-\frac{m}{n})}{n}} < p_0 < \frac{m}{n} + Z_\alpha \sqrt{\frac{\frac{m}{n}(1-\frac{m}{n})}{n}}\right\} = 1-\alpha$$

Dane

$m = 850$

$n = 1000$

$1 - \alpha = 0{,}95 \Rightarrow \alpha = 0{,}05$

$\Phi(z_\alpha) = 1 - \frac{\alpha}{2} = 0{,}975$

$z_\alpha = 1{,}96$

$$P\left\{0{,}85 - 1{,}96 \cdot \sqrt{\frac{0{,}85 \cdot 0{,}15}{1000}} < p_0 < 0{,}85 + 1{,}96\sqrt{\frac{0{,}85 \cdot 0{,}15}{1000}}\right\} = 0{,}95$$

$$P\{0{,}827 < p_0 < 0{,}872\} = 0{,}95$$

z prawdopodobieństwem 95% poziom ufności od 82,7% do 87,2 obejmuje odsetek osób, które przynajmniej mieli styczność z akcją reklamową.