



Politechnika Wrocławska

## Wydział Matematyki

Kierunek: matematyka stosowana

Specjalność: *nie dotyczy*

# Praca dyplomowa — inżynierska

## TESTOWANIE HIPOTEZ I ESTYMACJA W SYTUACJI POPULACJI SKOŃCZONEGO ROZMIARU

Kinga Kurowska

Słowa kluczowe:  
testowanie hipotez  
przedziały niepewności  
zastosowanie w medycynie

Krótkie streszczenie:

Tu znajdzie się krótki streszczenie mojej pracy... Tu znajdzie się krótki streszczenie mojej pracy... Tu znajdzie się krótki streszczenie mojej pracy... Tu znajdzie się krótki streszczenie mojej pracy... Tu znajdzie się krótki streszczenie mojej pracy... Tu znajdzie się krótki streszczenie mojej pracy... Tu znajdzie się krótki streszczenie mojej pracy...

Opiekun pracy	dr inż. Andrzej Giniewicz		
dyplomowej	Stopień naukowy, imię i nazwisko	Ocena	Podpis

Do celów archiwalnych pracę dyplomową zakwalifikowano do: \*

a) kategorii A (akta wieczyste),

b) kategorii BE 50 (po 50 latach podlegające ekspertyzie).

\* niepotrzebne skreślić

pieczęć wydziałowa

Wrocław, rok 2017



Politechnika Wrocławska

**Faculty of Pure and Applied Mathematics**

Field of study: Applied Mathematics

Specialty: *not applicable*

**Engineering Thesis**

**HYPOTHESIS TESTING AND ESTIMATION  
IN THE CASE OF FINITE POPULATION SIZE**

Kinga Kurowska

keywords:

hypothesis testing

tolerance and confidence regions

applications to biology and medical sciences

Short summary:

Here will be short summuary of my bachelor thesis... Here will be short summuary of my bachelor thesis... Here will be short summuary of my bachelor thesis... Here will be short summuary of my bachelor thesis... Here will be short summuary of my bachelor thesis...

Supervisor	dr inż. Andrzej Giniewicz		
	<i>Title, degree, name and surname</i>	<i>Grade</i>	<i>Signature</i>

*For the purposes of archival thesis qualified to: \**

*a) Category A (perpetual files)*

*b) Category BE 50 (subject to expertise after 50 years)*

*\* Delete as appropriate*

*stamp of the faculty*

Wrocław, 2017

# Spis treści

<b>Wstęp</b> . . . . .	5
<b>Rozdział 1. Schemat pobierania obserwacji</b> . . . . .	7
1.1. Nieskończona populacja . . . . .	7
1.2. Skończona populacja . . . . .	7
1.3. Porównanie rozkładów . . . . .	8
<b>Rozdział 2. Przedstawienie testów</b> . . . . .	11
2.1. Sformułowanie problemu . . . . .	11
2.2. Testy ze skończoną poprawką . . . . .	11
2.2.1. Test Z . . . . .	12
2.2.2. Test E . . . . .	12
2.2.3. Moc testu . . . . .	13
2.3. Test bez skończonej poprawki . . . . .	13
<b>Rozdział 3. Analiza testów</b> . . . . .	15
<b>Spis rysunków</b> . . . . .	19
<b>Bibliografia</b> . . . . .	21



# Wstęp

Początki teorii rachunku prawdopodobieństwa i statystyki sięgają XVI wieku. Na początku były to analizy rzutu kostką, czy prawdopodobieństwa błędów pomiarowych. Już w XVII Blaise Pascal sformułował i dowiódł własności trójkąta arytmetycznego oraz użył pojęcia kombinacji. Na początku XVIII wieku opublikowane zostały prace Jacoba Bernoulliego, w których zawarł wiele swoich tez na temat prawdopodobieństwa. Przez te kilka wieków teoria rachunku prawdopodobieństwa i statystyki znacząco się wzbogaciła i rozwinęła. Rozpoczęto rozważania na temat estymacji i testowania hipotez, które w naszych czasach, są zasadniczą domeną statystyki.

W przypadku dyskretnym najczęściej testowane są proporcje populacji. Chcemy się przekonać czy dana próbka ma jakąś konkretną proporcję, albo czy dwie próbki mają tę samą proporcję. Znana jest powszechnie teoria dotycząca testowania hipotez, gdy populacja jest nieskończona, a raczej na tyle duża, że możemy ją w przybliżeniu uznać za nieskończoną. Wtedy schemat próbkowania jest opisany przez rozkład Bernoulliego. Jednak przypadek nieskończonej populacji nie wyczerpuje tematu testowania proporcji. Gdy populacja jest bardzo mała albo, gdy próbka jest niewiele mniejsza od całej populacji, schemat próbkowania opiera się o rozkład hipergeometryczny.

Myślę, że warto zająć się teorią testowania hipotez dla skończonej populacji. W określonych przypadkach rozkład hipergeometryczny daje dużo dokładniejszą informację o badanym przypadku niż przybliżenie rozkładem dwumianowym. Ponadto zastosowanie tego typu testów ma duże znaczenie w medycynie, gdzie często rozważane populacje mają na tyle wyspecjalizowane cechy, że są uważane za małe. *Uzupełnić opis co się znajduje i gdzie!!!* W kolejnym rozdziale znajduje się opis dwóch testów opartych o rozkład hipergeometryczny oraz analiza prawdopodobieństwa błędu I rodzaju i mocy testu dla nich. Następnie wykonane jest porównanie tych testów z testem wykorzystującym rozkładem Bernoulliego.



## Rozdział 1

# Schemat pobierania obserwacji

W mojej pracy zajmuję się testami, które sprawdzają czy dwie populacje mają te same proporcje jakiejś badanej cechy. Zatem, aby wykonać taki test potrzebne są próbki z obu populacji. Wartość obserwacji to ilość osobników z badaną cechą w próbce, toteż proporcja jest liczona jako stosunek wartości do wielkości próbki. Zakładamy również, że populacje są od siebie niezależne, tym samym próbki pochodzące z tych populacji są także nie zależą od siebie.

### 1.1. Nieskończona populacja

Gdy populacja jest bardzo duża, możemy traktować ją jako nieskończoną. Wobec tego losowanie kolejnych elementów próbki jest niezależne, czyli jest to losowanie ze zwracaniem. Zakładamy, że pobierane obserwacje pochodzą z rozkładu Bernoullego  $\mathcal{B}(n, p)$ , gdzie  $n$  to rozmiar próby z nieskończonej populacji, a  $p$  to proporcja zdarzeń sprzyjających w populacji. Funkcja prawdopodobieństwa zmiennej losowej  $X$  z rozkładu dwumianowego jest równa

$$b(k; n, p) = P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad 0 \leq k \leq n. \quad (1.1)$$

### 1.2. Skończona populacja

Kiedy populację rozważamy jako skończoną kolejne elementy próbki są losowane bez zwracania. Oznacza to, że prawdopodobieństwo sukcesu zmienia się w trakcie pobierania elementów obserwacji w zależności od poprzednich. W takiej sytuacji próbka pochodzi z rozkładu hipergeometrycznego  $\mathcal{H}(n, M, N)$ . Przy czym  $n$  jest rozmiarem próbki,  $M$  ilością osobników w populacji z daną cechą, a  $N$  rozmiarem populacji. Zmienna losowa  $X$  z rozkładu hipergeometrycznego określa ma funkcja prawdopodobieństwa określona wzorem

$$h(k; n, M, N) = P(X = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}, \quad L \leq k \leq U, \quad (1.2)$$

gdzie  $L = \max\{0, M - N + n\}$  i  $U = \min\{n, M\}$ .

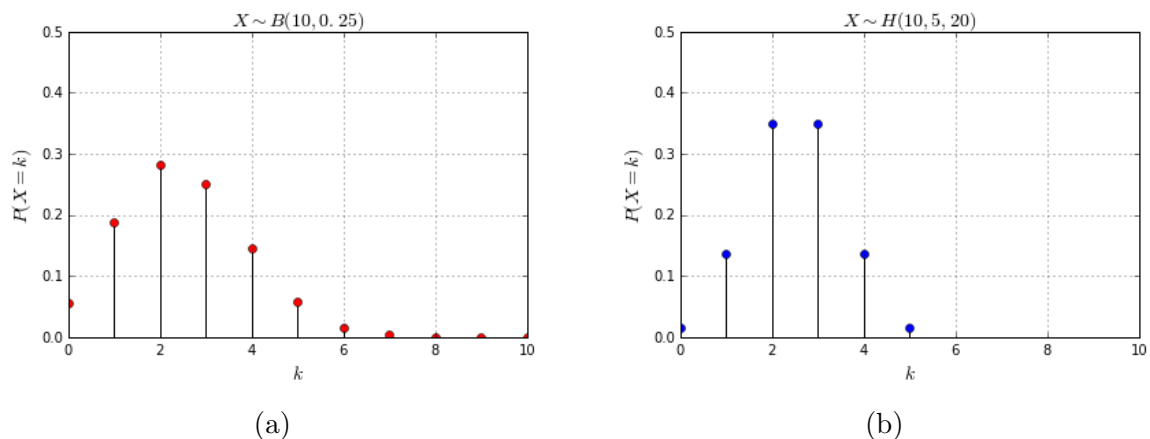
Zauważmy, że wzór funkcji  $h$  jest dość intuicyjny. Klasyczna definicja prawdopodobieństwa określa szansę zajścia zdarzenia  $A$  jako iloraz liczby zdarzeń elementarnych w  $A$  przez liczbę zdarzeń elementarnych w  $\Omega$ , czyli  $P(A) = |A|/|\Omega|$ . W tym przypadku zdarzeniem  $A$  jest to, że w próbce będzie  $k$  osobników z daną cechą. Zatem ilość zdarzeń elementarnych w  $A$  to kombinacje. Na jak wiele różnych sposobów możemy wybrać  $k$  osobników z  $M$  wszystkich posiadających daną cechę  $\binom{M}{k}$  razy możliwość wyborów

pozostałych osobników z reszty populacji  $\binom{N-M}{n-k}$ . Natomiast zbiór wszystkich zdarzeń elementarnych w  $\Omega$  to po prostu wybór losowej próbki  $n$  osobników z  $N$ -elementowej populacji  $\binom{N}{n}$ . Również ograniczenia nałożone na  $k$  są naturalne. Dolne ograniczenie  $L$  jest równe maksimum z 0 i  $M - N + n$ . Zatem będzie ono niezerowe, gdy  $M - N + n > 0$ . Przekształcając nierówność, otrzymujemy  $n > N - M$ . Taka postać jasno pokazuje, że jest to przypadek, w którym wielkość próbki przekracza ilość osobników w populacji bez badanej cechy. W konsekwencji czego mamy pewność, że w próbce będzie przynajmniej tyle osobników z daną cechą, ile wynosi różnica  $n - (N - M)$ . Ograniczenie górne jest mniej skomplikowane. Jest ono równe minimum z  $n$  i  $M$ , co jest oczywiste, ponieważ nie może być więcej osób w próbce z daną cechą niż w całej populacji. Cała ta analiza pokazuje, że rozkład hipergeometryczny jest ściśle związany z rozmiarem populacji.

### 1.3. Porównanie rozkładów

Rozkład hipergeometryczny daje bardzo podobne wyniki do rozkładu Bernoullego, gdy populacja jest duża, a próbka stosunkowo mała. Jednakże w przeciwnym wypadku widać znaczne różnice między tymi rozkładami. W szczególności obserwujemy, że zmienna losowa z rozkładu dwumianowego może osiągać wartości od 0 do  $n$ , ponieważ podczas brania próbki zakładamy losowanie ze zwracaniem. Natomiast w rozkładzie hipergeometrycznym kolejne losowania są od siebie zależne, prawdopodobieństwo wylosowania zmienia się w zależności od tego, co już wcześniej znalazło się w próbce.

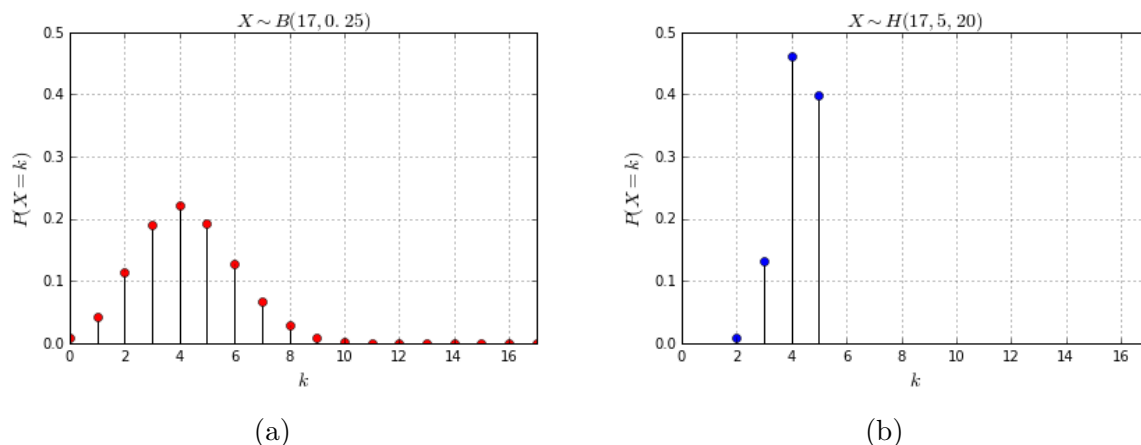
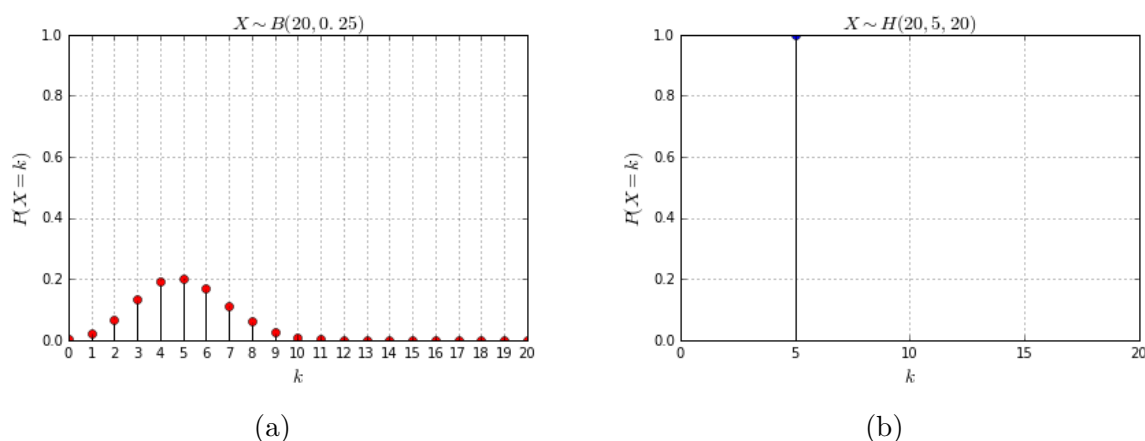
Rozważmy to na medycznym przykładzie. Załóżmy, że jest na świecie 20 osób, które są chore na jakąś bardzo rzadką chorobę oraz że 25% z nich ma szansę na wyzdrowienie. Chcemy dowiedzieć się, ile osób spośród przebadanych może wyzdrowieć. Weźmy 3 różne próbki o wielkościach  $n$  równych odpowiednio 10, 17, 20. Możemy tę sytuację opisać za pomocą rozkładem Bernoullego, wtedy badana zmienna losowa będzie z rozkładu  $\mathcal{B}(n, 0.25)$ . Drugim sposobem jest rozkład hipergeometryczny, wtedy zmienna losowa jest z rozkładu  $\mathcal{H}(n, 4, 20)$ . Na rysunkach 1.1 – 1.3 jest zobrazowana funkcja prawdopodobieństwa dla wymienionych przypadków.



Rysunek 1.1: Funkcja prawdopodobieństwa dla  $n = 10$

Im większa próbka tym widać większą różnicę między funkcjami obu rozkładów. Funkcja dla rozkładu hipergeometrycznego nie ma innych argumentów niż te, które są możliwe,



Rysunek 1.2: Funkcja prawdopodobieństwa dla  $n = 17$ Rysunek 1.3: Funkcja prawdopodobieństwa dla  $n = 20$ 

natomiast funkcja dla rozkładu dwumianowego jest liczona również dla nieprawdopodobnych argumentów, ponadto wtedy jej wartość jest niezerowa. Bardzo skrajny przypadek przedstawia rysunek 1.3, próbka równa się populacji, czyli tak naprawdę wiemy już wszystko. Wykres 1.3b idealnie obrazuje nam ten przypadek. Pewne jest to, że w 20 osobach będzie dokładnie 5, które mogą wyzdrowieć. Natomiast wykres 1.3a kompletnie nie pokazuje tego. Głównie ze względu na już wspomniane uwzględnianie nieprawdopodobnych argumentów, przez co prawdopodobieństwo rozkłada się na pozostałe przypadki. Warto również zaznaczyć, że im większa próbka, tym prawdopodobieństwo, że  $X = 5$  dla rozkładu Bernoulliego jest coraz mniejsze, ponieważ mamy więcej elementów w obserwacji. Tymczasem dla rozkładu hipergeometrycznego jest wręcz odwrotnie, to prawdopodobieństwo rośnie, aż w końcu osiąga wartość 1. Co jest dużo bardziej logiczne, bo im więcej przebadaliśmy osobników, tym więcej wiemy o próbce i jest bardziej prawdopodobne, że znajdzie się w niej aż 5 szczęśliwych pacjentów.

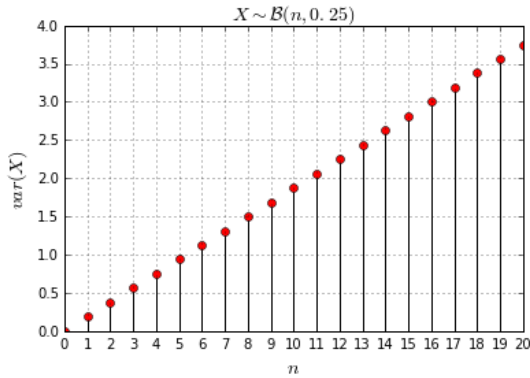
Spójrzmy jeszcze, jak wyglądają średnia i wariancja dla rozważanych rozkładów. Gdy  $X \sim \mathcal{B}(n, p)$

$$E(X) = np, \quad \text{Var}(X) = np(1 - p). \quad (1.3)$$

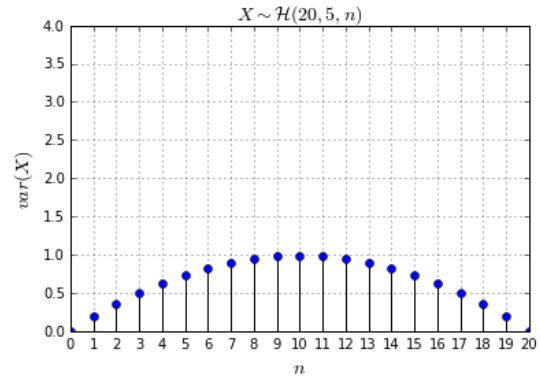
Podczas gdy  $X \sim \mathcal{H}(n, M, N)$

$$E(X) = n \frac{M}{N}, \quad \text{Var}(X) = n \frac{M}{N} \left(1 - \frac{M}{N}\right) \frac{N-n}{N-1}. \quad (1.4)$$

Parametr  $p$  rozkładu dwumianowego odpowiada ilorazowi parametrów  $M/N$  w rozkładzie hipergeometrycznym. Wobec tego średnie obu rozkładów są sobie równe, a wariancje różni dodatkowy składnik w rozkładzie hipergeometrycznym  $(N-n)/(N-1)$ . Przeanalizujemy jak ten czynnik wpływa na zróżnicowanie rozkładów. Rysunek 1.4 przedstawia wykresy wariancji dla rozważanego przykładu w zależności od wielkości obserwacji. W przypadku rozkładu dwumianowego wariancja stale rośnie wraz ze wzrostem próbki. Ostatecznie, gdy już przebadamy wszystkich chorych pacjentów, ma ona największą wartość. Myśląc zdroworozsądkowo, nie jest to poprawny wynik. Oczekujemy raczej, że gdy przebadamy już wszystkich, wariancja osiągnie wartość 0, ponieważ nie ma wtedy losowości. Taki rezultat daje nam wykres 1.4b. Funkcja na początku rośnie, ale gdy wielkość próbki przekroczy połowę rozmiaru populacji, wariancja zaczyna maleć aż do zera. Odzwierciedla to fakt, że gdy coraz więcej wiemy o populacji, losowość uzyskanych wyników maleje.



(a)



(b)

Rysunek 1.4: Wariancja w zależności od rozmiaru próbki

## Rozdział 2

# Przedstawienie testów

W niniejszym rozdziale są przedstawione trzy testy, jeden oparty na rozkładzie dwumianowym i dwa wykorzystujące rozkład hipergeometryczny. Jest także omówiony sposób liczenia dla nich mocy testu.

### 2.1. Sformułowanie problemu

Założmy, że  $X_1$  i  $X_2$  są niezależnymi zmiennymi losowymi. Zaobserwowane wartości  $X_1$  i  $X_2$  oznaczmy odpowiednio  $k_1$  i  $k_2$  oraz proporcje w obserwacjach  $p_1$  i  $p_2$ . Będę zajmować się testowaniem hipotez

$$H_0: p_1 = p_2 \quad \text{przeciwko} \quad H_1: p_1 \neq p_2, \quad (2.1)$$

w oparciu o wartości obserwacji i znane parametry populacji. Rozważmy unormowaną statystykę

$$Z_{X_1, X_2} = \frac{X_1/n_1 - X_2/n_2}{\sqrt{V_{X_1, X_2}}}, \quad (2.2)$$

gdzie  $V_{X_1, X_2}$  to estymator wariancji rozkładu łącznego zmiennej losowej  $X_1/n_1 - X_2/n_2$  pod warunkiem prawdziwości  $H_0$ . Jego wzór zależy od rozkładu, z którego pochodzą zmienne losowe  $X_1$  i  $X_2$ . Wartość statystyki  $Z_{X_1, X_2}$  oznaczmy jako  $Z_{k_1, k_2}$ . Jest ona wyliczana według powyższych wzorów, poprzez zamienienie  $X_1$  i  $X_2$  odpowiednio wartościami obserwacji  $k_1$  i  $k_2$ .

### 2.2. Testy ze skończoną poprawką

Jak już było wspomniane w sekcji 1.2, próbki w przypadku skończonej populacji pochodzą z rozkładu hipergeometrycznego. Zatem  $X_1 \sim \mathcal{H}(n_1, M_1, N_1)$ ,  $X_2 \sim \mathcal{H}(n_2, M_2, N_2)$  oraz proporcje są równe  $p_1 = M_1/N_1$ ,  $p_2 = M_2/N_2$ . Znane parametry to rozmiary próbek  $n_1$  i  $n_2$  oraz wielkości populacji  $N_1$  i  $N_2$ .

W celu wyprowadzenia wariancji rozkładu łącznego  $X_1/n_1 - X_2/n_2$  pod warunkiem  $p_1 = p_2$ , zapiszmy łączną wariancję rozważanej zmiennej losowej, korzystając z własności wariancji oraz tego, że  $Cov(X_1, X_2) = 0$  z niezależności  $X_1$  i  $X_2$

$$Var(X_1/n_1 - X_2/n_2) = Var(X_1/n_1) + Var(X_2/n_2) = Var(X_1)/n_1^2 + Var(X_2)/n_2^2. \quad (2.3)$$

Wiemy, że wariancje  $X_1$  i  $X_2$  są równe

$$Var(X_1) = n_1 p_1 (1 - p_1) (N_1 - n_1) / (N_1 - 1), \quad (2.4)$$

$$Var(X_2) = n_2 p_2 (1 - p_2) (N_2 - n_2) / (N_2 - 1). \quad (2.5)$$

Pamiętając, że zakładamy równość  $p_1 = p_2$  zastąpmy oba parametry jednym równym  $p$ . Po podstawieniu otrzymujemy

$$\begin{aligned} V_{X_1, X_2} &= \frac{1}{n_1} p(1-p) \frac{N_1 - n_1}{N_1 - 1} + \frac{1}{n_2} p(1-p) \frac{N_2 - n_2}{N_2 - 1} = \\ &= p(1-p) \left( \frac{N_1 - n_1}{n_1(N_1 - 1)} + \frac{N_2 - n_2}{n_2(N_2 - 1)} \right), \end{aligned} \quad (2.6)$$

przy czym  $p$  to proporcja liczby osobników z daną cechą do całości populacji w rozkładzie łącznym. Zatem jest ona równa  $p = (X_1 + X_2)/(n_1 + n_2)$ . Ostatecznie otrzymujemy

$$V_{X_1, X_2} = \left( \frac{N_1 - n_1}{n_1(N_1 - 1)} + \frac{N_2 - n_2}{n_2(N_2 - 1)} \right) \left( \frac{X_1 + X_2}{n_1 + n_2} \right) \left( 1 - \frac{X_1 + X_2}{n_1 + n_2} \right). \quad (2.7)$$

### 2.2.1. Test Z

Test Z jest oparty na centralnym twierdzeniu granicznym, które mówi, że pod warunkiem prawdziwości  $H_0$  rozważana statystyka  $Z_{X_1, X_2}$  jest w przybliżeniu z rozkładu standardowego normalnego  $\mathcal{N}(0, 1)$ . Wtedy  $p$ -wartość wyraża się wzorem

$$P(|Z_{X_1, X_2}| \geq |Z_{k_1, k_2}| | H_0) = 2(1 - \Phi(|Z_{k_1, k_2}|)), \quad (2.8)$$

gdzie  $\Phi$  oznacza dystrybuantę rozkładu  $N(0, 1)$ . Test Z odrzuca hipotezę zerową, gdy  $p$ -wartość jest mniejsza od poziomu istotności  $\alpha$ .

### 2.2.2. Test E

Test E opiera się o rzeczywistą  $p$ -wartość, która, według artykułu K. Krishnamoorthy i J. Thomson z 2002 roku, jest równa [1]

$$\begin{aligned} P(|Z_{X_1, X_2}| \geq |Z_{k_1, k_2}| | H_0) &= E_{X_1, X_2}(\mathbf{1}(|Z_{X_1, X_2}| \geq |Z_{k_1, k_2}|) | H_0) = \\ &= \sum_{x_1=L_1}^{U_1} \sum_{x_2=L_2}^{U_2} h(x_1; n_1, N_1 p, N_1) h(x_2; n_2, N_2 p, N_2) \mathbf{1}(|Z_{X_1, X_2}| \geq |Z_{k_1, k_2}|), \end{aligned} \quad (2.9)$$

gdzie  $E_{X_1, X_2}$  to wartość oczekiwana łącznego rozkładu  $(X_1, X_2)$ , a  $p$  jest nieznaną wspólną proporcją pod warunkiem  $H_0$ . Nie jest możliwe policzenie  $p$ -wartości wprost ze wzoru (2.9), ponieważ nie znamy parametru proporcji  $p$ . Krishnamoorthy i Thomson (2002) zaproponowali estymator  $p$ -wartości [1]

$$\begin{aligned} P(|Z_{X_1, X_2}| \geq |Z_{k_1, k_2}| | H_0) &= \\ &= \sum_{x_1=L_{x_1}}^{U_{x_1}} \sum_{x_2=L_{x_2}}^{U_{x_2}} h(x_1; n_1, \hat{M}_1, N_1) h(x_2; n_2, \hat{M}_2, N_2) \mathbf{1}(|Z_{X_1, X_2}| \geq |Z_{k_1, k_2}|), \end{aligned} \quad (2.10)$$

przy czym  $\hat{p} = (k_1 + k_2)/(n_1 + n_2)$ ,  $\hat{M}_i = [N_i \hat{p}]$  oraz  $L_{x_i} = \max\{0, \hat{M}_i - N_i + n_i\}$ ,  $U_{x_i} = \min\{n_i, \hat{M}_i\}$ ,  $i = 1, 2$ . Test odrzuca  $H_0$  wtedy, gdy  $p$ -wartość wyliczona wg wzoru (2.10) jest mniejsza od poziomu istotności  $\alpha$ .

### 2.2.3. Moc testu

Moc testu to prawdopodobieństwo odrzucenia hipotezy zerowej, gdy jest ona nieprawdziwa. Toteż jest ona wyznacznikiem dobrego testu. Większa wartość mocy oznacza lepszy test.

Moc obu testów można wyliczyć, korzystając z funkcji prawdopodobieństwa rozkładu hipergeometrycznego. Dla testu Z pod warunkiem hipotezy alternatywnej  $H_1$  moc, zgodnie z rozważaniami Krishnamoorthy i Thomson (2002), jest równa [1]

$$\sum_{k_1=L_1}^{U_1} \sum_{k_2=L_2}^{U_2} h(k_1; n_1, M_1, N_1) h(k_2; n_2, M_2, N_2) \mathbb{1}(|Z_{k_1, k_2}| > z_{1-\alpha/2}), \quad (2.11)$$

gdzie  $L_i = \max\{0, M_i - N_i + n_i\}$  i  $U_i = \min\{n_i, M_i\}$ , a  $z_{1-\alpha/2}$  oznacza kwantyl rozkładu normalnego standardowego rzędu  $1 - \alpha/2$ .

Tymczasem dla testu E, wg Krishnamoorthy i Thomson (2002), moc zdefiniowana jest następująco [1]

$$\begin{aligned} & \sum_{k_1=L_1}^{U_1} \sum_{k_2=L_2}^{U_2} h(k_1; n_1, M_1, N_1) h(k_2; n_2, M_2, N_2) \times \\ & \times \mathbb{1} \left( \sum_{x_1=L_{x_1}}^{U_{x_1}} \sum_{x_2=L_{x_2}}^{U_{x_2}} h(x_1; n_1, \hat{M}_1, N_1) h(x_2; n_2, \hat{M}_2, N_2) \mathbb{1}(|Z_{X_1, X_2}| \geq |Z_{k_1, k_2}|) \leq \alpha \right), \end{aligned} \quad (2.12)$$

gdzie wszelkie parametry oznaczają się takie same jak we wzorach (2.10) i (2.11).

## 2.3. Test bez skończonej poprawki

Dla testu bez poprawki na skończony rozmiar populacji, zamiast rozkładu hipergeometrycznego używamy dwumianowego, a więc  $X_1$  i  $X_2$  są niezależnymi zmiennymi losowymi o rozkładzie Bernoulliego  $X_1 \sim \mathcal{B}(n_1, p_1)$ ,  $X_2 \sim \mathcal{B}(n_2, p_2)$ . Znane parametry to rozmiary próbek  $n_1$  i  $n_2$ .

Wariancję rozkładu łącznego  $X_1/n_1 - X_2/n_2$  pod warunkiem  $p_1 = p_2$  możemy wyprowadzić analogicznie jak w sekcji 2.2, wychodząc od łącznej wariancji rozważanej zmiennej losowej

$$\text{Var}(X_1/n_1 - X_2/n_2) = \text{Var}(X_1)/n_1^2 + \text{Var}(X_2)/n_2^2. \quad (2.13)$$

Wariancje  $X_1$  i  $X_2$  są równe

$$\text{Var}(X_1) = n_1 p_1 (1 - p_1), \quad (2.14)$$

$$\text{Var}(X_2) = n_2 p_2 (1 - p_2). \quad (2.15)$$

Zastępując  $p_1$  i  $p_2$  jednym parametrem równym  $p$ , otrzymujemy

$$V_{X_1, X_2} = p(1 - p)/n_1 + p(1 - p)/n_2 = p(1 - p)(1/n_1 + 1/n_2) \quad (2.16)$$

przy czym  $p = (X_1 + X_2)/(n_1 + n_2)$ .

Ten test jest, podobnie jak omówiony wcześniej test E, oparty o estymator  $p$ -wartości, który, zgodnie z artykułem Storer i Kim z 1990 roku, jest równy [2]

$$P(|Z_{X_1, X_2}| \geq |Z_{k_1, k_2}| | H_0) = \sum_{x_1=0}^{n_1} \sum_{x_2=0}^{n_2} b(x_1; n_1, \hat{p}_1) b(x_2; n_2, \hat{p}_2) \mathbb{1}(|Z_{X_1, X_2}| \geq |Z_{k_1, k_2}|), \quad (2.17)$$

gdzie  $\hat{p} = (k_1 + k_2)/(n_1 + n_2)$ . Test odrzuca hipotezę zerową, gdy  $p$ -wartość jest mniejsza od poziomu istotności  $\alpha$ .

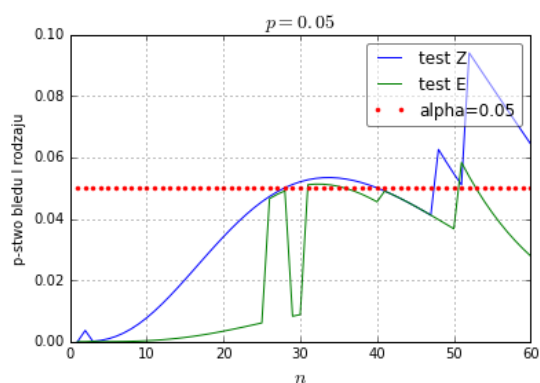
Moc testu jest również podobna do mocy testu E. W tym przypadku wyraża się wzorem

$$\sum_{k_1=0}^n \sum_{k_2=0}^n b(k_1; n_1, p_1) b(k_2; n_2, p_2) \times \\ \times \mathbb{1} \left( \sum_{x_1=0}^{n_1} \sum_{x_2=0}^{n_2} b(x_1; n_1, \hat{p}_1) b(x_2; n_2, \hat{p}_2) \mathbb{1} (|Z_{X_1, X_2}| \geq |Z_{k_1, k_2}|) \leq \alpha \right). \quad (2.18)$$

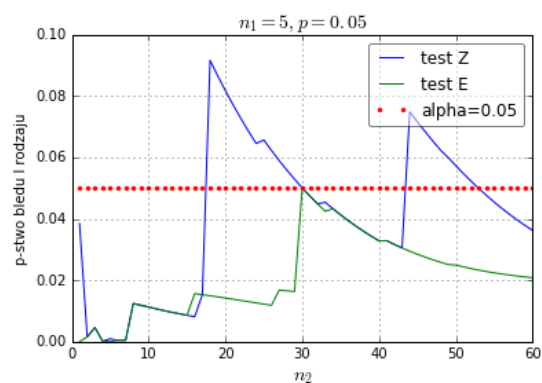
## Rozdział 3

# Analiza testów

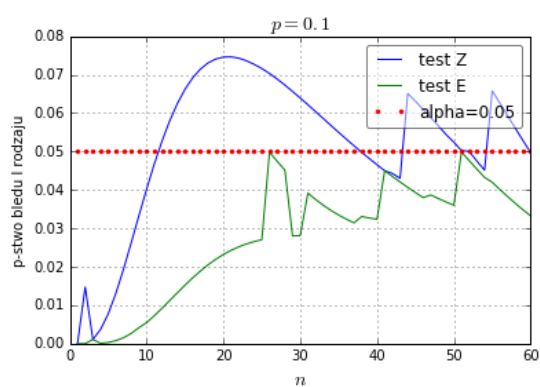
Do porównania testów napisałam programy, które wyliczają prawdopodobieństwo błędu I rodzaju oraz moc testu dla różnych parametrów.



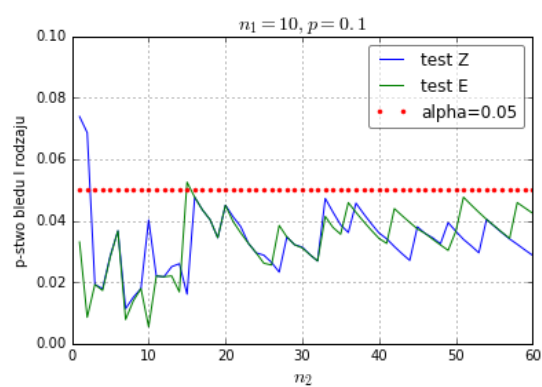
(a)



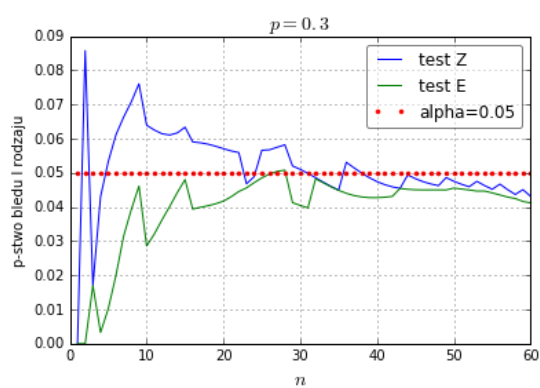
(b)



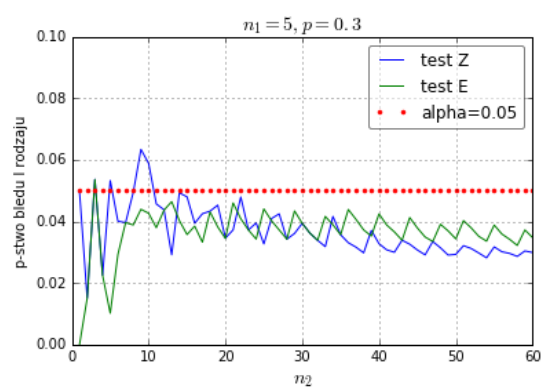
(c)



(d)



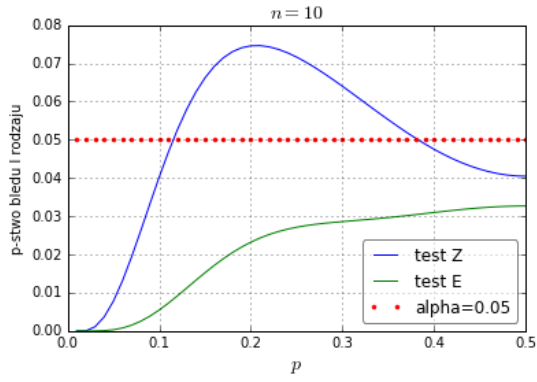
(e)



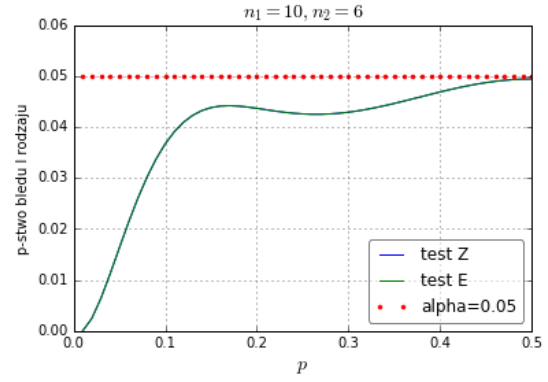
(f)

Rysunek 3.1: P-stwo błędu I rodzaju jako funkcja rozmiaru próbki  $n$ , przy zadanym poziomie istotności  $\alpha = 0.05$ ;  $N_1 = N_2 = 100$

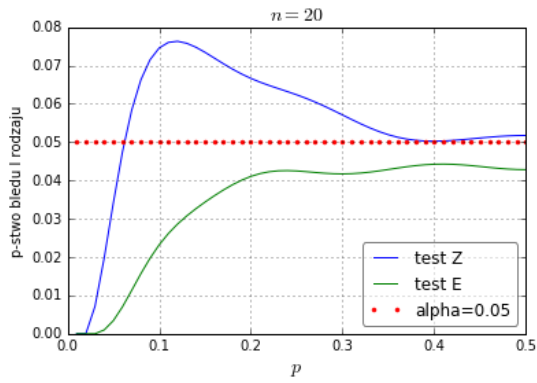




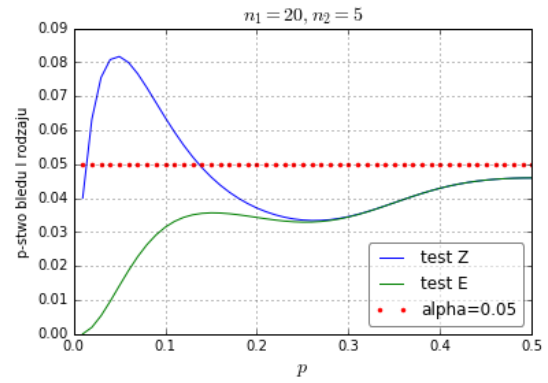
(a)



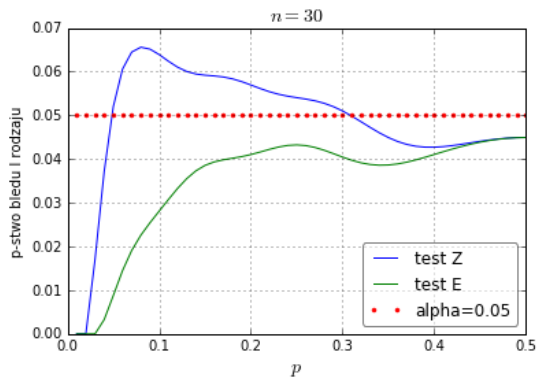
(b)



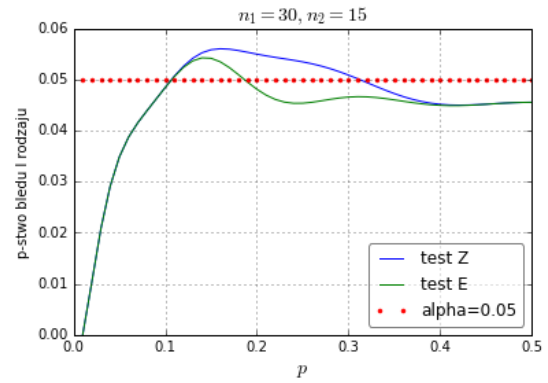
(c)



(d)

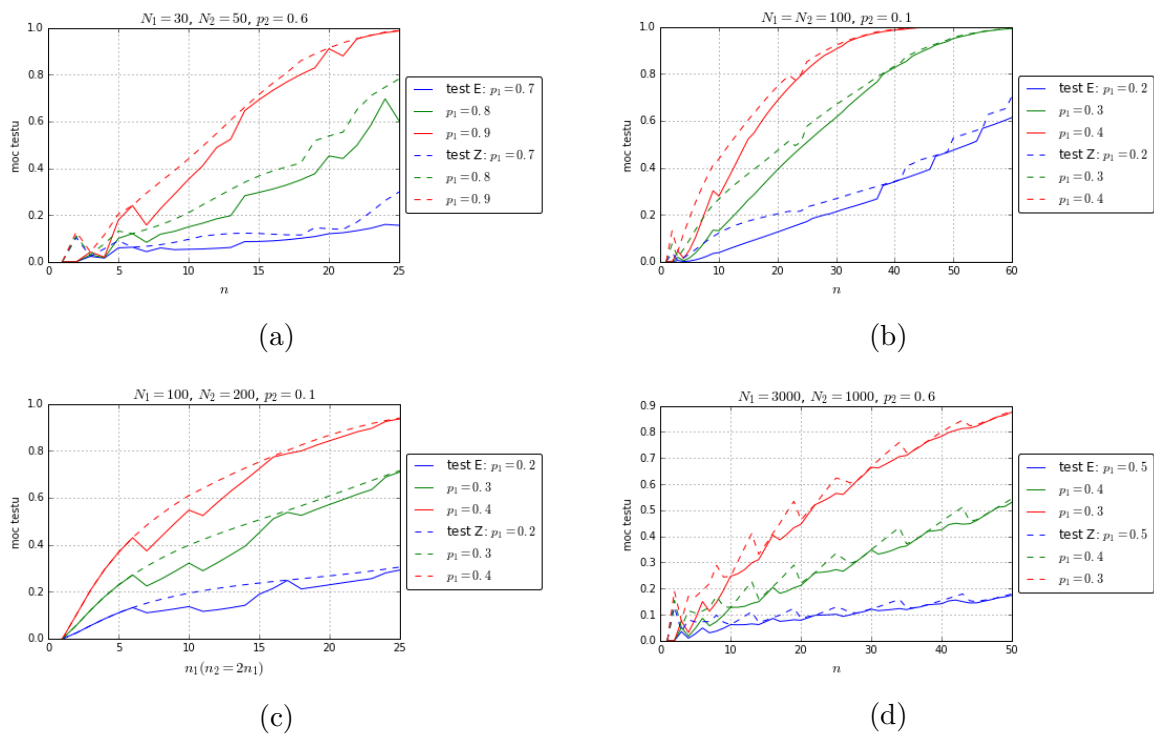


(e)



(f)

Rysunek 3.2: P-stwo błędu I rodzaju jako funkcja parametru  $p = M_1/N_1 = M_2/N_2$ , przy zadanym poziomie istotności  $\alpha = 0.05$ ;  $N_1 = N_2 = 100$



Rysunek 3.3: Moc testu jako funkcja rozmiaru próbki  $n$ , przy zadanym poziomie istotności  $\alpha = 0.05$

# Spis rysunków

1.1	Funkcja prawdopodobieństwa dla $n = 10$ . . . . .	8
1.2	Funkcja prawdopodobieństwa dla $n = 17$ . . . . .	9
1.3	Funkcja prawdopodobieństwa dla $n = 20$ . . . . .	9
1.4	Wariancja w zależności od rozmiaru próbki . . . . .	10
3.1	P-stwo błędu I rodzaju jako funkcja rozmiaru próbki $n$ , przy zadanym poziomie istotności $\alpha = 0.05$ ; $N_1 = N_2 = 100$ . . . . .	16
3.2	P-stwo błędu I rodzaju jako funkcja parametru $p = M_1/N_1 = M_2/N_2$ , przy zadanym poziomie istotności $\alpha = 0.05$ ; $N_1 = N_2 = 100$ . . . . .	17
3.3	Moc testu jako funkcja rozmiaru próbki $n$ , przy zadanym poziomie istotności $\alpha = 0.05$	18



# Bibliografia

- [1] K. Krishnamoorthy and Jessica Thomson. Hypothesis testing about proportions in two finite populations. *The American Statistician*, 56(1):215–222, 2002.
- [2] Barry E. Storer and Choongrak Kim. Exact properties of some exact test statistics for comparing two binomial proportions. *Journal of the American Statistical Association*, 85:146–155, 1990.