ĆWICZENIE STATYSTYCZNE. STABILIZACJA WZGLĘDNYCH CZĘSTOŚCI I ROZKŁADY WZGLĘDNYCH CZĘSTOŚCI ZDARZEŃ

I. Cel ćwiczenia: zapoznanie studenta z charakterystykami i prawidłowościami zdarzeń statystycznych na przykładzie prostych doświadczeń. Przeprowadzenie ćwiczenia statystycznego ma zilustrować w sposób przekonujący sens wprowadzenia niektórych pojęć i wykazać, że zjawiskami masowymi rządzą określone prawidłowości, które często można przewidzieć teoretycznie w oparciu o dość proste założenia.

II. Przyrządy: przyrząd - ruletka, stoper.

III. Literatura:

1. Z. Hellwig, Elementy rachunku prawdopodobieństwa i statystyki matematycznej, W-wa 1978, PWN.
2. Z. Pawłowski, Wstęp do statystyki matematycznej, W-wa 1966, PWN.
3. J.L. Kacperski, I pracownia fizyczna, UŁ 1982.
4. R. Zieliński, Tablice statystyczne, W-wa 1972, PWN.

Wiele zjawisk, nie tylko fizycznych, ma charakter przypadkowy, tzn. wynik zjawiska nie może być przewidziany przed jego zakończeniem ze względu na niemożność uwzględnienia zbyt wielu warunkujących go czynników lub wręcz ich nieznajomość.

Jakkolwiek nie można przewidzieć wyniku pojedynczego zdarzenia, to jednak wielokrotne powtarzanie prób wywołania tego zdarzenia pozwala dla „przeciętnych" wyników otrzymać pewne prawidłowości zwane statystycznymi

IV.1 Stabilizacja częstości zdarzeń i pojęcie prawdopodobieństwa.

Elementarnym pojęciem w dziedzinie zjawisk statystycznych jest ***zdarzenie.*** Zdarzeniem jest np. wyrzucenie orła w przypadkowym rzucie monetą lub wyrzucenie trójki w przypadkowym rzucie kostką do gry. Zdarzeniem jest także otrzymanie wyniku pomiaru zawierającego się w z góry określonym przedziale wartości.

Rzut monetą, kostką lub przeprowadzenie pomiaru nazywa się próbą. Najprostszą klasyfikacją dla zdarzeń jest stwierdzenie, czy w wyniku próby określone zdarzenie nastąpiło czy też nie. Jeżeli zdarzenie nastąpiło, to mówimy o sukcesie i takiemu zdarzeniu przyporządkowujemy liczbę ***1***, jeśli zaś nie nastąpiło, to mówimy o niepowodzeniu i przyporządkowujemy mu liczbę ***0***.

Niech liczba wszystkich prób w doświadczeniu będzie równa n, zaś liczbę sukcesów dla określonego rodzaju zdarzeń oznaczmy przez m. Stosunek ***m/n*** będziemy nazywali ***"częstością względną"*** występowania interesującego nas zjawiska.

Na ogół doświadczenie potwierdza przypuszczenie, że częstość względna w miarę wzrostu liczby prób n, mimo pewnych wahań, wykazuje wyraźną tendencję zdążania do określonej dla danego zdarzenia granicy zwanej jego prawdopodobieństwem, co zapisujemy

$$\lim_{n\to\infty} \frac{m}{n} = P \ . \qquad (1)$$

Oznacza to, że różnica między częstością względną m/n a prawdopodobieństwem P mimo lokalnych wahań staje się coraz mniejsza przy wzroście n. Ta prawidłowość nosi nazwę stabilizacji względnej częstości zdarzeń.

Zdarzenie, dla którego P = 0, nazywać będziemy niemożliwym, a takie dla którego P = 1, pewnym.

IV.2 Prawdopodobieństwo zdarzenia, wyznaczanie jego wartości i rozkład względnych częstości

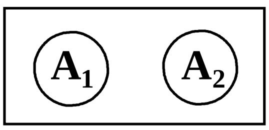
Z poprzedniego wynika, że prawdopodobieństwo jest wielkością charakteryzującą pewną prawidłowość statystyczną zjawiska przypadkowego. Z sensu jaki nadaje prawdopodobieństwu wzór (1) należy wnioskować, że wartość prawdopodobieństwa zawiera się w granicach $0 \le P \le 1$, ponieważ $0 \le m \le n$ (tj. liczba sukcesów nie może być większa od liczby prób) oraz nie może przyjmować wartości ujemnych.

Jednym ze sposobów znajdowania przybliżonej wartości P podaje wzór (1) po wykonaniu dużej liczby prób. Słabą stroną tego sposobu jest to, że praktycznie nie można przeprowadzić nieskończonej liczby prób, a w związku z tym należy przyjąć P ≈ m/n, co daje wartość przybliżoną a ponadto wartość prawdopodobieństwa może być wyznaczona po przeprowadzeniu doświadczenia (serii prób) i nie może być znana przed jego wykonaniem tj. a priori.

Niekiedy istnieje możliwość wyznaczenia prawdopodobieństwa a priori. Konieczne jest wówczas, by interesujące nas zdarzenie stanowiło równoprawną część spośród innych wariantów zdarzeń. Np. przy rzucaniu monetą mamy dwie możliwości: wyrzucenia orła i wyrzucenia resztki, zaś przy rzucaniu kostką do gry - sześć wariantów zdarzeń: wyrzucenie jedynki, dwójki itd. aż do szóstki. Jeśli moneta i kostka są symetryczne względem środka ciężkości, to każdy z możliwych wyników jest równoprawny. W pierwszym przypadku zarówno wyrzucenie orła jak i resztki stanowią 1/2 część wszystkich wariantów, a w przypadku kostki wyrzucenie wybranej liczby oczek (w granicach 1 - 6) stanowi 1/6 część wszystkich możliwych zdarzeń.

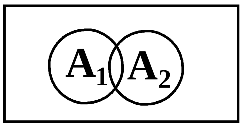
Wartości stanowiące odwrotność liczby wszystkich równoprawnych wariantów zdarzeń równe są prawdopodobieństwu zaistnienia jednego wariantu.

IV.3 Zdarzenia niezależne, wykluczające się, przeciwstawne, iloczyn, suma zdarzeń i ich prawdopodobieństwa.



Rys.1

O dwóch zdarzeniach A_1 i A_2 powiemy, że są ***niezależne*** wtedy, gdy zajście lub niezaistnienie jednego z nich nie zmienia prawdopodobieństwa realizacji drugiego zdarzenia. Np. zdarzenie polegające na wylosowaniu kuli określonej barwy z urny zawierającej np. dwa rodzaje kul w sytuacji, gdy w poprzednim losowaniu dokonano zmiany składu urny w sposób umowny zależnie od wyniku tego losowania, jest zdarzeniem zależnym od rezultatu poprzedniego ciągnienia.



Rys.2

Jeżeli w ciągu zdarzeń A_1 , A_2 , ...A_n żadne dwa spośród nich nie mogą zaistnieć łącznie, to zdarzenia te nazywamy wzajemnie ***wykluczającymi się.*** Interpretując zdarzenie losowe jako zbiór utworzony z elementów zbioru zawierającego wszystkie możliwe zdarzenia (elementy), o zdarzeniach wykluczających się powiemy, że ich zbiory są rozłączne (czyli nie zawierają elementów wspólnych). Sytuację tą ilustruje rys.1.

Ponieważ zdarzenie polegające na łącznym zaistnieniu zdarzeń A_1 i A_2 nazywa się ***iloczynem zdarzeń***, to prawdopodobieństwo iloczynu zdarzeń wykluczających się jest równe zeru. Iloczyn zdarzeń ilustruje rys.2. Prawdopodobieństwo iloczynu zdarzeń niewykluczających się jest równe

$$P(A_1 \cdot A_2) = P(A_1) \cdot P(A_2)\,, \qquad (2)$$

gdzie $P(A_1)$ i $P(A_2)$ są bezwzględnymi prawdopodobieństwami odpowiednich zdarzeń A_1 i A_2 .

Jako przykład iloczynu dwóch zdarzeń można podać zdarzenie polegające na wyciągnięciu dwóch kul białych przy dwóch ciągnieniach z tej samej urny lub z różnych urn.

Sumą zdarzeń A_1 , A_2 , ... A_n jest zdarzenie polegające na zrealizowaniu się przynajmniej jednego z tych zdarzeń. Np. jeśli A_1 jest zdarzeniem polegającym na wyrzuceniu orła w rzucie monetą, natomiast A_2 jest zdarzeniem polegającym na wyciągnięciu losu pustego, to suma zdarzeń $A_1 + A_2$ oznacza zdarzenie, w którym albo wynikiem rzutu monetą będzie orzeł, albo zostanie wyciągnięty los pusty, albo że zostanie wyciągnięty los pusty w ciągnieniu z urny i wynikiem rzutu monetą będzie wyrzucenie orła. Zauważmy, że sumą zdarzeń będzie również zdarzenie polegające na wyrzuceniu parzystej liczby oczek przy rzucie kostką. Zdarzenie to zrealizuje się zawsze wtedy, gdy wynikiem rzutu będzie albo dwójka, albo czwórka, albo szóstka.

Jeżeli zdarzenia parami wykluczają się, to prawdopodobieństwo sumy zdarzeń jest równe sumie prawdopodobieństw realizacji każdego z tych zdarzeń, co zapisujemy

$$P(A_1 + A_2) = P(A_1) + P(A_2)\,. \qquad (3)$$

Jeżeli natomiast zdarzenia nie wykluczają się tzn. poza alternatywnym zaistnieniem obu zdarzeń nie jest wykluczone łączne ich zajście, wówczas prawdopodobieństwo sumy tych zdarzeń będzie pomniejszone w stosunku do (3) o prawdopodobieństwo ich iloczynu

$$P(A_1 + A_2) = P(A_1) + P(A_2) - P(A_1 \cdot A_2)\,. \qquad (4)$$

Np. wyniki rzutu jedną monetą są niezależne i wykluczają się. Natomiast wyniki doświadczenia polegającego na rzucie dwiema monetami są niezależne ale określony wynik dla jednej monety nie wyklucza identycznego wyniku dla drugiej.

Zdarzenie losowe polegające na niewystąpieniu zdarzenia losowego A nazywamy ***zdarzeniem przeciwstawnym*** do A. Jeżeli prawdopodobieństwo realizacji zdarzenia A jest P(A) a zdarzenie przeciwstawne oznaczymy przez $\overline{A}$, to prawdopodobieństwo tego, że A nie wystąpi jest równe

$$P(\overline{A}) = 1 - P(A)\,, \qquad (5)$$

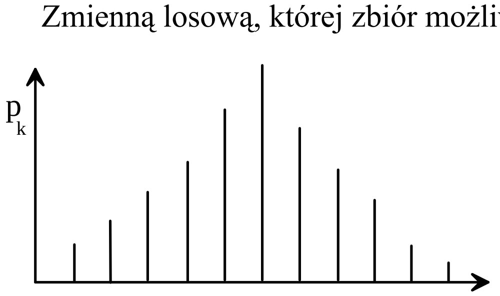
tzn. że prawdopodobieństwo zdarzenia losowego i prawdopodobieństwo zdarzenia przeciwstawnego uzupełniają się do jedności.

IV.4 Zmienna losowa

Zmienną przyjmującą w sposób losowy (przypadkowy) wartości z określonego przedziału nazywamy ***zmienną losową.***

W przypadku doświadczenia polegającego na rzutach kostką do gry powiemy, że dziedzina zmiennej losowej jest skończona tzn. zmienna ta przyjmuje wartości 1, 2, ... 6 z określonymi prawdopodobieństwami równymi odpowiednio prawdopodobieństwu wyrzucenia jedynki, dwójki itd.

Zmienną losową, której zbiór możliwych wartości jest skończony lub co najwyżej przeliczalny nazywamy zmienną losową dyskretną. Jeżeli natomiast zmienna losowa może przyjąć każdą wartość rzeczywistą z pewnego przedziału skończonego lub nieskończonego, to taką zmienną nazwiemy zmienną losową ciągłą.



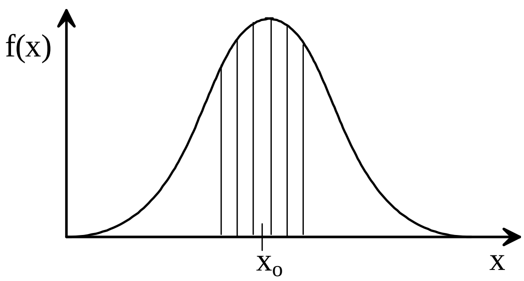
Rys.3

IV.5 Funkcje rozkładu zmiennych losowych.

Zależność wiążąca wartości, jakie może przyjmować zmienna losowa dyskretna z prawdopodobieństwem ich występowania, nazywa się funkcją prawdopodobieństwa

(rys.3). W przypadku ciągłej zmiennej losowej mówimy o funkcji gęstości prawdopodobieństwa f(x) (rys.4). Prawdopodobieństwo przyjęcia przez zmienną losową ciągłą wartości z przedziału (x_o-$\Delta x/2$, x_o+ $\Delta x/2$) o szerokości Δx na tyle małej, by można przyjąć, że w nim f(x)= const., jest równe

$$P(x_o - \Delta x/2 < x < x_o + \Delta x/2) = f(x) \cdot \Delta x \,. \tag{6}$$



Rys.4

Dla Δx→0 szukane elementarne prawdopodobieństwo wynosi

$$dP = f(x)dx \,. \tag{7}$$

W przypadku, gdy przedział Δx jest na tyle rozciągły, że f(x) nie pozostaje w nim stała, obliczenie prawdopodobieństwa wymaga policzenia całki

$$P(x_o - \Delta x/2 < x < x_o + \Delta x/2) = \int_{x_o - \Delta x/2}^{x_o + \Delta x/2} f(x)dx \,. \tag{8}$$

Należy zauważyć, że nie każda funkcja zmiennej losowej może być funkcją rozkładu prawdopodobieństwa. Musi ona spełniać tzw. warunek normalizacyjny. Dla zmiennej ciągłej jest on następujący:

$$\int_{-\infty}^{\infty} f(x)dx = 1 \,, \tag{9}$$

a dla zmiennej dyskretnej:

$$\sum_{i=1}^{n} P_i = 1 \ . \tag{10}$$

Czasem zachodzi potrzeba scharakteryzowania rozkładu zmiennej losowej za pomocą jednej lub paru liczb wyrażających najistotniejsze własności rozkładu. Do takich charakterystyk rozkładu zalicza się ***nadzieję matematyczną***, inaczej wartość oczekiwaną zmiennej losowej i ***wariancję***.

Nadzieja matematyczna informuje o tym, jaki jest przeciętny poziom wartości przybieranych przez zmienną losową. Dla zmiennej losowej dyskretnej nadzieję matematyczną określa się jako sumę iloczynów poszczególnych wartości jakie zmienna może przyjąć przez odpowiadające im prawdopodobieństwa

$$E(x) = \sum_k x_k \cdot P(x = x_k) \ . \tag{11}$$

Sumowanie rozciąga się na wszystkie możliwe wartości zmiennej losowej.

Jeżeli zmienna losowa jest ciągła i przybiera wartości jedynie z przedziału [a, b], to jej nadzieja matematyczna wyraża się całką

$$E(x) = \int_a^b x \cdot f(x)dx \ . \tag{12}$$

Wariancję definiuje się jako nadzieję matematyczną kwadratów odchyleń zmiennej losowej x od jej wartości oczekiwanej. Zgodnie z podanym określeniem zapiszemy

$$D(x) = E[x - E(x)]^2 \,. \tag{13}$$

W obliczeniach ma zastosowanie inna postać wzoru (13), a mianowicie

$$D(x) = E(x^2) - E^2(x) . \qquad (14)$$

Na ogół większe znaczenie praktyczne ma pierwiastek kwadratowy z wariancji, który określa się mianem odchylenia standardowego zmiennej losowej. Oznaczając odchylenie standardowe przez σ mamy

$$\sigma = \sqrt{E[x - E(x)]^2} \ . \qquad (15)$$

IV.6 Rozkłady zmiennych losowych

IV.6.1 *Rozkład dwupunktowy*

Rozkład, w którym dyskretna zmienna losowa może przyjmować tylko dwie wartości x_1 i x_2 z prawdopodobieństwem odpowiednio p i q (p + q = 1) nazywamy rozkładem dwupunktowym. Zmienna losowa o rozkładzie dwupunktowym interpretuje się jako liczbę przypisaną realizacji określonego zdarzenia w pojedynczym doświadczeniu. Na ogół realizacji zdarzenia losowego przypisuje się liczbę 1, a nie zaistnieniu tego zdarzenia - liczbę 0.
Prawo rozkładu dwupunktowego ma postać

$$\begin{cases} P(x=1) = p \\ P(x=0) = q \quad = 1 - p \end{cases}$$

Nadzieja matematyczna zmiennej o takim rozkładzie wynosi

$$E(x) = p , \qquad (16)$$

natomiast wariancja ma wartość

$$D(x) = p \cdot q . \qquad (17)$$

IV.6.2. Rozkład dwumianowy

Rozkład dwumianowy jest rozkładem zmiennej losowej dyskretnej zdefiniowanej jako liczba realizacji określonego wyniku w n niezależnych próbach, gdy w każdej z nich prawdopodobieństwo pojawienia się tego wyniku jest p.

Prawo rozkładu zmiennej losowej o rozkładzie dwumianowym ma postać

$$P(x = m) = \binom{n}{m} p^m (1-p)^{n-m} \ , \qquad (18)$$

gdzie

$$\binom{n}{m} = \frac{n!}{m!(n-m)!} . \qquad (19)$$

Nadzieja matematyczna dla tego rozkładu wynosi

$$E(x) = n \cdot p \ , \qquad (20)$$

a wariancja

$$D(x) = n \cdot p \cdot q . \qquad (21)$$

Dla przypadku p = q rozkład dwumianowy jest symetryczny względem m = n/2.

IV.6.3. Rozkład Poissona

Przy spełnieniu warunku, że prawdopodobieństwo otrzymania interesującego wyniku w pojedynczej próbie jest bardzo małe oraz liczba prób dąży do nieskończoności przy czym jednocześnie zachodzi

$$n \cdot p = \lambda = const \text{ (i niezbyt duże),} \tag{22}$$

rozkład dwumianowy dąży do rozkładu Poissona:

$$P(x = m) = \frac{\lambda^m}{m!} e^{-\lambda}, \tag{23}$$

λ jest średnią liczbą realizacji wyróżnionego zdarzenia w serii n doświadczeń.

Dla rozkładu Poissona nadzieja matematyczna i wariancja są sobie równe

$$E(x) = D(x) = \lambda . \tag{24}$$

IV.6.4. Rozkład prostokątny

Najprostszym rozkładem ciągłym jest rozkład prostokątny, kiedy to zmienna losowa z jednakowym prawdopodobieństwem może przyjmować każdą wartość z przedziału [a, b]. Gęstość prawdopodobieństwa tej zmiennej losowej w przedziale [a, b] jest stała.

$$f(x) = c = const\ . \tag{25}$$

Uwzględniając warunek normalizacyjny (9) znajdziemy, że

$$c = \frac{1}{b - a}\ . \tag{26}$$

Zatem funkcję rozkładu prostokątnego zapiszemy następująco

$$f(x) = \begin{cases} 0 & x < a \\ 1/(b - a) & a \le x \le b \\ 0 & x > b \end{cases} . \tag{27}$$

Nadzieja matematyczna $$E(x) = \frac{b + a}{2}\ . \tag{28}$$

Wariancja $$D(x) = \frac{(b - a)^2}{12}\ . \tag{29}$$

IV.6.5 Rozkład normalny

Szczególnie ważną rolę w statystyce matematycznej pełni rozkład normalny o gęstości prawdopodobieństwa

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x - \mu)^2}{2\sigma^2} \right] \tag{30}$$

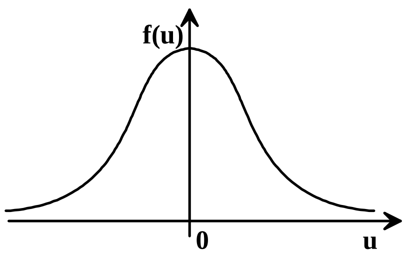
μ, σ są parametrami rozkładu.

Nadzieja matematyczna $$E(x) = \mu\ . \tag{31}$$

Wariancja $$D(x) = \sigma^2\ . \tag{32}$$

W zastosowaniach rozkładu normalnego wykorzystuje się zmienną losową u związaną ze zmienną x następującą zależnością:

$$u = \frac{x - \mu}{\sigma}\ , \tag{33}$$



Rys.5 Funkcja gęstości prawdopodobieństwa rozkładu normalnego N(0,1)

Można wykazać, że zmienna losowa u ma rozkład normalny o parametrach $\mu = 0$ i $\sigma = 1$ [N(0,1)]. Funkcję gęstości tego rozkładu określa formuła

$$f(u) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{u^2}{2}\right] \qquad (34)$$

Gęstość prawdopodobieństwa dla rozkładu normalnego z $\mu = 0$ i $\sigma = 1$ zostały ze względów praktycznych stabelaryzowane. Stabelaryzowano również dystrybuantę tego rozkładu czyli funkcję określającą prawdopodobieństwo przyjęcia przez zmienną losową normalną wartości mniejszej od górnej granicy całkowania

$$F(x) = \int_{-\infty}^{x} f(x)dx \ . \qquad (35)$$

IV.6.6 Rozkład χ^2 (chi-kwadrat)

Jeśli x_1, x_2, ..., x_j jest ciągiem niezależnych zmiennych losowych o rozkładzie normalnym z $\mu = 0$ i $\sigma = 1$, to zmienną losową:

$$\chi^2 = x_1^2 + x_2^2 + \cdots + x_j^2 \qquad (36)$$

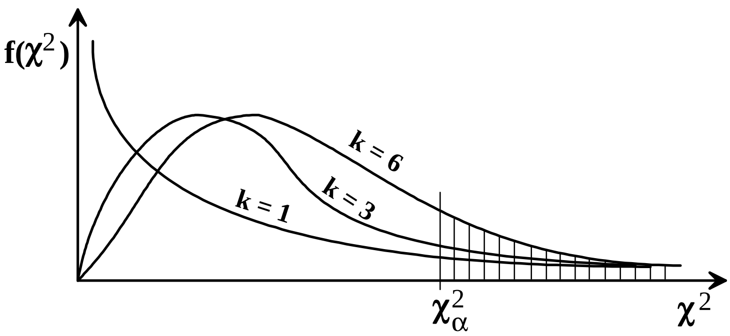
nazywamy zmienną o rozkładzie χ^2 z k stopniami swobody(k jest liczbą niezależnych składników w wyrażeniu (36)).

Funkcja gęstości prawdopodobieństwa zmiennej losowej o rozkładzie χ^2 i k stopniach swobody dana jest wzorem

$$f(\chi^2) = \frac{1}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})}(\chi^2)^{\frac{k}{2}-1}\exp(-\frac{\chi^2}{2}) \ . \qquad (37)$$

Nadzieja matematyczna $E(\chi^2) = k,$ (38)

Wariancja $D(\chi^2) = 2k.$ (39)



Rys.6
Przebieg funkcji gęstości prawdopodobieństwa zmiennej losowej χ^2 z liczbą stopni swobody odpowiednio: k = 1, k = 3, k = 6.

Funkcja χ^2 została stabelaryzowana ze względu m. in. na jej znaczenie w teorii weryfikacji hipotez statystycznych. Istnieją dwa rodzaje tablic. Jeden podaje dla różnych wartości parametru k (a więc dla różnych rozkładów χ^2) prawdopodobieństwo tego, że zmienna losowa przyjmuje wartość większą od określonej liczby χ^2_α

$$p = \alpha = P(\chi^2 > \chi^2_\alpha) \qquad (40)$$

Drugi rodzaj tablic podaje dla różnych wartości parametru k takie liczby rzeczywiste χ^2_α, że prawdopodobieństwo przybrania przez zmienną losową wartości większej od danej liczby jest równe z góry danej liczbie α

$$\alpha = P(\chi^2 > \chi^2_\alpha) \qquad (41)$$

Parametr α nazywamy ***poziomem istotności***. Jest on równy polu powierzchni między daną krzywą a osią χ^2 w zakreskowanym obszarze na rysunku 6.

IV.7 Test χ^2 (chi-kwadrat)

Test χ^2 ma zastosowanie przy sprawdzaniu hipotezy, że rozkład obserwowanej zmiennej losowej ma określoną postać analityczną.

Załóżmy, że przedmiotem obserwacji jest pewna zmienna losowa x i że należy sprawdzić hipotezę, iż zmienna ta ma określony typ rozkładu.

Podzielmy obszar zmienności x na j odrębnych klas (przedziałów). Oznaczmy przez n_i liczbę zaobserwowanych w próbie elementów, dla których zmienna x przyjęła wartość należącą do klasy i. Niech prawdopodobieństwo tego, że zmienna losowa przyjmie wartość należącą do tej klasy (wynikające z założonej hipotezy H_o) wynosi P_i . Jeżeli przez n oznaczymy liczbę obserwacji w próbie, to iloczyn $n_i^o = n \cdot P_i$ nazwiemy liczebnością teoretyczną i-tej klasy. Jest to oczekiwana liczba obserwacji w i-tej klasie obliczona przy założeniu, że prawdziwa jest hipoteza sprawdzana H_o.

Różnica między liczbami n_i oraz n_i^o może być podstawą do zbudowania sprawdzianu hipotezy, że obserwowana zmienna losowa ma określony typ rozkładu. Gdy różnice są małe, będziemy skłonni przyjąć hipotezę, natomiast przy dużych rozbieżnościach między rzeczywistymi a teoretycznymi liczebnościami hipotezę odrzucamy.

Jako sprawdzian hipotezy przyjmujemy (za K. Pearsonem) wyrażenie

$$\chi^2 = \sum_{i=1}^{j} \frac{(n_i - n_i^o)^2}{n_i^o} \; . \qquad (42)$$

Można wykazać, że przy $n \to \infty$ wyrażenie po prawej stronie wzoru (42) ma rozkład χ^2 , stąd oznaczenie lewej strony powyższej zależności tym symbolem. Rozkład ten ma $k = j - r - 1$ stopni swobody, przy czym r oznacza tu liczbę parametrów rozkładu o założonej postaci funkcyjnej, które należało wstępnie oszacować z próby, by następnie obliczyć prawdopodobieństwa P_i i liczebności teoretyczne $n_i^o = n \cdot P_i$.

Z testu tego można skorzystać tylko wtedy, gdy wszystkie liczby n_i są równe przynajmniej 5, a n jest przynajmniej równe 50.

Tak więc testowanie rozkładu będzie polegało na obliczeniu wartości χ^2 zgodnie ze wzorem (42) i porównaniu jej ze stabelaryzowaną zmienną $\chi^2_{k;\alpha}$ (patrz np. odpowiednia tabela *I pracownia fizyczna J.L. Kacperski)*. Hipotezę zgodności rozkładu eksperymentalnego i teoretycznego przyjmujemy wówczas, gdy spełniona jest nierówność:

$$\chi^2 \le \chi^2_{k;\alpha} \qquad (43)$$

gdzie α jest przyjętym poziomem istotności np. $\alpha = 0{,}05$, k - liczbą stopni swobody.

V Wykonanie pomiarów.

Wykonanie pomiarów statystycznych sprowadza się do przeprowadzenia ok. 200 ruchów ruletką i zanotowania poniższych danych:

1. wielkości kąta sektora $\Delta\beta_j$, na którym zatrzymała się strzałka (np. dla $\Delta\beta$ = 20°, 40°, 60°, 80°, 160°).

2. położenie wskazówki na skali kątowej (kąt β_i).

3. czas "t" trwania ruchu strzałki od momentu wprawienia jej w ruch do chwili zatrzymania.

Dane te najwygodniej jest zapisać w tabeli:

Lp.	$\Delta\beta_i$	β_i	t
1			
n			

VI. Opracowanie wyników.

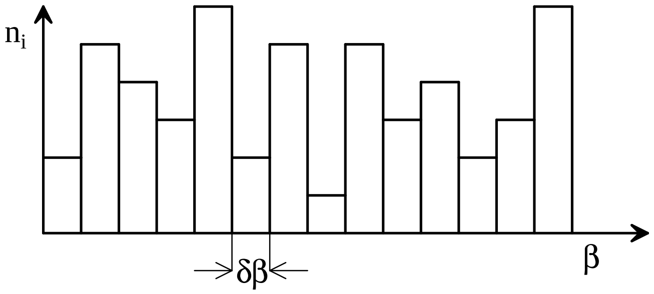
W opracowaniu otrzymanych danych statystycznych należy:

1a. Dla wybranego sektora o wielkości $\Delta\beta$ wykreślić zależność stosunku m_i/n od n dla wartości n wzrastających np. o 5 (m_i oznacza ilość przypadków zatrzymania się strzałki w wybranym sektorze w serii n prób). Przeanalizować przebieg krzywej łączącej kolejne punkty wykresu z punktu widzenia stabilizacji względnych częstości.

1b. Wykonać rozkład częstości zatrzymania się strzałki m razy w serii np. n = 9 niezależnych prób w wybranym sektorze $\Delta\beta = 60^o$, gdzie m =1, 2, 3, ... n. Porównać otrzymany rozkład zmiennej dyskretnej z rozkładem dwumianowym, dla którego prawdopodobieństwo zajścia wyróżnionego zdarzenia wynosi p = 60/360 = 0,166.

2a. Wykonać rozkład częstości dla wielkości kąta β_i (zmiennej losowej ciągłej), wskazywanego przez strzałkę. Zbudować histogram doświadczalny dzieląc cały zakres zmienności kąta [0, 360^o] na klasy (przedziały) o ustalonej szerokości $\Delta x = \delta\beta$.

Szerokość $\delta\beta$ należy przyjąć jako równą w przybliżeniu 1/4 odchylenia standardowego σ postulowanego rozkładu (w tym przypadku prostokątnego). Ponieważ teoretyczna wartość σ dla tego konkretnego rozkładu prostokątnego wynosi ok.104^o (patrz wzór (29) i (15)), to proponowana szerokość przedziału wynosi $\delta\beta \approx (1/4)\cdot 104^o = 26^o$. Z uwagi na to, że ilość przedziałów musi być liczbą całkowitą, ostatecznie bierzemy $\delta\beta = 24^o$ i wówczas ilość przedziałów jest równa j = 360^o/24^o = 15. Na osi Y odłożyć liczebności n_i zatrzymań strzałki w poszczególnych klasach.



$$\sum_{i=1}^{j} n_i = n$$

Obliczyć:

wartość średnią
$$\bar{\beta} = \frac{\sum_{i=1}^{n} \beta_i}{n},$$

średni błąd kwadratowy
$$s_\beta = \sqrt{\frac{\sum_{i=1}^{n}\left(\beta_i - \bar{\beta}\right)^2}{n-1}}.$$

2b. Sprawdzić hipotezę , iż rozkład $n_i(\beta)$ jest rozkładem prostokątnym.
Prawdopodobieństwo tego, że zmienna losowa czyli kąt β przyjmie wartość należącą do i-tego przedziału wynosi

$$P_i = \frac{1}{b-a}\delta\beta = \frac{\delta\beta}{360^o}$$

a liczebność teoretyczna $n_i^o = n \cdot P_i = n\frac{\delta\beta}{360^o}$.

Na tle rozkładu doświadczalnego narysować rozkład teoretyczny i zaznaczyć parametry rozkładu μ i σ (patrz wzór (28) oraz (29) i (15)).

Ze wzoru (42) znajdujemy χ^2 . Obliczenia zebrać w tabeli:

Nr przedziału	n_i	$n_i^o = n\frac{\delta\beta}{360^o}$	n_i - n_i^o	$\frac{(n_i - n_i^o)^2}{n_i^o}$

Pamiętajmy, by n_i i n_i^o w przedziałach były większe od 5. W przeciwnym razie przedziały należy połączyć, by $n_i \geq 5$. Liczba stopni swobody jest równa liczbie słupków histogramu po połączeniu (jeśli to było konieczne) minus 2 (dla rozkładu prostokątnego r = 1). Przy przyjętym poziomie istotności α i dla obliczonej liczby stopni swobody k odczytać z tablic (np. I pracownia fizyczna J. L. Kacperski) wartość $\chi^2_{k,\alpha}$. Porównać tę wartość z otrzymaną i ustosunkować się do postawionej hipotezy.

3. Wykonać rozkład doświadczalny dla czasu "t" trwania ruchu strzałki oraz wyliczyć rozkład teoretyczny zakładając określoną funkcję rozkładu prawdopodobieństwa np. rozkładu normalnego. Zakładając rozkład normalny należy obliczyć:

wartość średnią

$$\bar{t} = \frac{\sum_{i=1}^{n} \cdot t_i}{n} ,$$

średni błąd kwadratowy

$$s_t = \sqrt{\frac{\sum_{i=1}^{n} (t_i - \bar{t})^2}{n-1}} .$$

Przedstawić wyniki doświadczalne w formie graficznej w postaci histogramu doświadczalnego, dzieląc przedział zmienności zmiennej losowej t na j klas o szerokości Δt każda (wg podobnych zasad jak w przypadku histogramu prostokątnego).

Wyniki doświadczalne będące podstawą histogramu zapisać w tabeli

Nr przedziału	Środek przedz. t_i	Liczba przypadków n_i	$t_i - \bar{t}$
1			
...			
j			

Rozkład doświadczalny normalizujemy za pomocą podstawienia $u_i = \frac{t_i - \bar{t}}{s_t}$ (nowa zmienna losowa). Z tablic standaryzowanego rozkładu normalnego odczytujemy gęstość prawdopodobieństwa dla u_i czyli $p(u_i)$. Wyrażenie

$$P_i = p(u_i) \cdot \Delta u_i = p(u_i) \cdot \frac{a}{s_t}$$

jest przybliżoną wartością prawdopodobieństwa otrzymania wyniku w i-tym przedziale o szerokości *Δu,* gdzie *a* = *Δt* jest szerokością słupka histogramu dla zmiennej "t" . Precyzja przybliżenia zależy od szerokości Δt słupka histogramu.
Na tym samym wykresie co histogram doświadczalny zaznaczyć punkty rozkładu teoretycznego łącząc je ze sobą linią ciągłą (wykorzystać czwartą kolumnę w tabeli poniżej). Zaznaczyć parametry tego rozkładu tzn. punkty: $t = \bar{t}$, $t = \bar{t} \pm s_t$ wraz z rzędnymi. Obliczyć wartość χ^2 i w sposób podobny jak w punkcie **2b** ocenić czy założenie o normalności rozkładu było słuszne. Wyniki można zebrać w tabeli:

Nr. przedziału	$u_i = \frac{t_i - \bar{t}}{s_t}$	$p(u_i)$	$n_i^o = n \cdot \frac{\Delta t}{s_t} \cdot p(u_i)$	$n_i - n_i^o$	$\frac{(n_i - n_i^o)^2}{n_i^o}$
1					
...					
j					
					$\chi^2 =$

Czwarta kolumna w tabeli podaje wartości oczekiwane n_i^o w i-tym przedziale.
Do opracowania danych doświadczalnych o zakładanym rozkładzie normalnym może posłużyć również program mikrokomputerowy HISTPAS.EXE dostępny w Pracowni Mikrokomputerowej i w I Pracowni Fizycznej.

UWAGA

Obszerniejszy zarys teorii zagadnień przedstawionych w tej instrukcji można znaleźć w tekście instrukcji pod takim samym tytułem oznaczonej jako* M-18BIS *dostępnej w bibliotece Fizyki.