SIGMA KWADRAT

LUBELSKI KONKURS STATYSTYCZNO-DEMOGRAFICZNY

Weryfikacja hipotez statystycznych

WNIOSKOWANIE STATYSTYCZNE

Wnioskowanie statystyczne, to proces uogólniania wyników uzyskanych na podstawie próby na całą populację generalną.

Podstawą uogólniania wyników uzyskanych na podstawie próby na całą populację generalną jest tzw. *statystyczna próba losowa* będąca podzbiorem populacji generalnej, który podlega bezpośredniej obserwacji statystycznej ze względu na interesujący nas rozkład cechy.

Na podstawie wyników (cech) obserwowanych dla próby można obliczyć np. takie charakterystyki rozkładu jak: średnia arytmetyczna czy odchylenie standardowe.
Te charakterystyki nazywane są parametrami z próby lub statystykami.

Statystyka to dowolna charakterystyka rozkładu obliczona na podstawie wyników próby losowej (może być opisana za pomocą zapisu funkcyjnego f(x)).

Statystyka jest zmienną losową i posiada swój rozkład, który zależy od:

- schematu losowania,
- liczebności próby,
- rozkładu zbiorowości, z której wylosowano próbę.

SPOSOBY POBIERANIA PRÓBY LOSOWEJ

Schemat losowania próby, to praktyczny sposób pobierania elementów populacji generalnej do próby.

Wyróżnić można na przykład losowanie:

- zależne i niezależne (ze zwracaniem),
- warstwowe (proporcjonalne, optymalne),
- indywidualne i zespołowe,
- jednostopniowe i wielostopniowe,
- ograniczone i nieograniczone.

Do wyboru próby można wykorzystać np. następujące techniki losowania:

- *tablice liczb losowych* – to zbiory 2, 3 , 4 cyfrowych liczb o zupełnie przypadkowej kolejności cyfr (określonej w niezależnym losowaniu),
- *generatory liczb losowych* – to specjalne programy komputerowe generujące liczby losowe.

SPOSOBY POBIERANIA PRÓBY LOSOWEJ

Losowanie niezależne polega na zachowaniu tych samych warunków podczas losowania kolejnych jednostek do próby. W wyniku takiego losowania każda wylosowana jednostka jest ponownie włączana do populacji generalnej, zatem ta sama jednostka może być wylosowana nawet kilkukrotnie (losowanie niezależne jest typowe dla populacji nieskończonych, losowanie zależne – jednostka raz wylosowana nie bierze udziału w dalszym losowaniu - jest stosowane dla populacji skończonych). *W przypadku licznej zbiorowości losowania zależne i niezależne dają praktycznie takie same wyniki*.

W losowaniu indywidualnym losujemy pojedyncze elementy populacji, a w zespołowym (grupowym) losowaną jednostką jest pewien zespół badanych jednostek (próbę tworzą wszystkie elementy z wylosowanych grup).

Losowanie jednostopniowe cechuje się jednym stadium losowania (elementy populacji są od razu wylosowywane – ostatecznie). Losowanie nazywamy wielostopniowym, gdy występuje kilka stopni losowania.

Losowanie nieograniczone odbywa się się z całej populacji generalnej, a ograniczone dokonuje się z poszczególnych części populacji np. losowanie warstwowe.

Losowanie warstwowe polega na tym, że całą zbiorowość generalną dzieli się na jednorodne wewnętrznie części, a następnie z każdej z nich dokonuje się losowania.

Aby można było odnieść wyniki badania próby do zbiorowości generalnej
PRÓBA MUSI BYĆ REPREZENTATYWNA.

Reprezentatywność próby zależy od sposobu wyboru próby (losowy, celowy) oraz liczebności próby.

Można mówić, że próba jest reprezentatywna, jeżeli spełnia warunki:

1) elementy populacji pobierane są do próby w sposób losowy,
2) próba jest dostatecznie liczna.

Jeżeli próba została wybrana w sposób losowy i jest dostatecznie liczna, to mówimy, że próba jest reprezentatywna. Oznacza to, że z dużym prawdopodobieństwem można sądzić, iż struktura próby będzie zbliżona do struktury zbiorowości.

Inaczej mówiąc, aby wyniki badania można było uogólnić na całą populację generalną struktura próbki musi być możliwie najbardziej zbliżona do struktury badanej populacji (musi być ona miniaturą populacji generalnej) i mieć określoną liczebność (dostateczną).

Rozkład cechy w populacji generalnej nazywamy rozkładem teoretycznym, natomiast rozkład cechy w próbie nazywamy rozkładem empirycznym.

REPREZENTATYWNOŚĆ PRÓBY

Przyjmując, że rozkład próby jest rozkładem empirycznym, a rozkład populacji jest rozkładem teoretycznym, to wzajemną relację pomiędzy tymi rozkładami można wyrazić za pomocą dystrybuanty empirycznej i dystrybuanty teoretycznej (twierdzenie Gliwienki).

Oznaczenie:

$\sup\limits_{-\infty<x<\infty} |G(x) - F(x)|$ - kres górny bezwzględnej różnicy między dystrybuantą empiryczną a dystrybuantą teoretyczną.

Tw. : Niech F(x) będzie dystrybuantą teoretyczną, a G(x) dystrybuantą empiryczną w n-elementowej próbie. Jeżeli wyniki losowania elementów do próby są zdarzeniami niezależnymi, to:

$$P\left\{ \sup_{-\infty<x<\infty} |G(x) - F(x)| \xrightarrow[n\to\infty]{} 0 \right\} = 1$$

Wniosek:

Gdy próba losowa jest dostatecznie liczna, to można uważać z prawdopodobieństwem bliskim jedności, że rozkład empiryczny mało różni się od rozkładu teoretycznego. Inaczej mówiąc próba jest tym bardziej reprezentatywna, im jest liczniejsza a wnioskowanie jest bardziej precyzyjne.

ROZKŁADY Z PRÓBY

Ze względu na liczebność próby rozkłady statystyk możemy podzielić na rozkłady:

- dokładne i
- graniczne.

Rozkładem dokładnym statystyki U nazywamy jej rozkład prawdopodobieństwa wyznaczony dla każdej liczby naturalnej *n*. Rozkłady te dotyczą badań reprezentacyjnych, dla których liczba obserwacji n jest mała. Rozkłady dokładne nazywane są również rozkładami z małych prób.

Rozkładem granicznym statystyki U nazywamy taki rozkład prawdopodobieństwa tej statystyki, do którego przybliża się jej rozkład przy $n \to \infty$.

!!! Nie ma jednej liczby n od której można uznać daną próbę za dużą. Zależy to od szybkości zbieżności rozkładu danej statystyki U do jej rozkładu granicznego. Dla niektórych statystyk rozkład jest dokładny już przy n > 30, ale niekiedy dopiero dla n > 100 daje dobre przybliżenie rozkładu statystyki z jej rozkładem granicznym.

W badaniach statystycznych rolę rozkładu granicznego spełnia *rozkład normalny*. Natomiast *rozkłady dokładne* są różne w zależności od tego, jaką statystyką U posługujemy się w procesie wnioskowania statystycznego.

ELEMENTY PROCESU WNIOSKOWANIA STATYSTYCZNEGO

ESTYMACJA PARAMETRÓW STRUKTURALNYCH ZBIOROWOŚCI GENERALNEJ

Estymacją nazywamy szacowanie wartości parametrów lub rozkładu zmiennej losowej w populacji generalnej na podstawie rozkładu empirycznego, uzyskanego z próby losowej pobranej z tej populacji.

Wyróżnia się estymację parametrów punktową oraz przedziałową.

Można również mówić o estymacji parametrycznej i nieparametrycznej.

Estymacja punktowa - polega na wybraniu jednej liczby, która na podstawie wyników uzyskanych z próby ma być najlepszym oszacowaniem danego parametru w populacji generalnej. Precyzję tego estymatora mierzy średni błąd szacunku estymatora, którym jest odchylenie standardowe błędów szacunku jakie popełnia się szacując parametr Q na podstawie wielu n-elementowych prób.

Estymacja przedziałowa - polega na konstrukcji na podstawie wyników uzyskanych w próbie pewnego przedziału liczbowego, który z prawdopodobieństwem bliskim jedności (z góry zadanym) zawierałby prawdziwą wartość parametru w populacji generalnej. Zadane z góry prawdopodobieństwo nazywane jest współczynnikiem ufności i oznacza się $1 - \alpha$, a cały wyznaczony (oszacowany) przedział nazywa się przedziałem ufności.

ESTYMACJA PRZEDZIAŁOWA PARAMETRÓW

Interpretacja poziomu ufności (współczynnika ufności) jest następująca: przy wielokrotnym pobieraniu prób n-elementowych z populacji i wyznaczaniu na ich podstawie granic przedziałów ufności, średnio w (1 - α) · 100% przypadków otrzymujemy przedziały pokrywające nieznaną wartość szacowanego parametru.

Poziom ufności (1 - α) jest bliski jedności i zwykle przyjmuje się: 0,90, 0,95, 0,98, 0,99.

Sposób konstrukcji przedziału ufności związany jest z rozkładem odpowiedniego estymatora. Rozkład ten jest zależny z kolei od założeń dotyczących rozkładu cechy w zbiorowości generalnej oraz od liczebności próby.

PRZEDZIAŁ UFNOŚCI DLA WARTOŚCI PRZECIĘTNEJ
DUŻA PRÓBA

Jeżeli liczebność próby jest duża (n > 30) to konstrukcję przedziału ufności opiera się na statystyce:

$$u = \frac{\bar{x} - m}{\sigma(x)} \sqrt{\mathrm{n}} \quad \text{lub} \quad u = \frac{\bar{x} - m}{s(x)}$$

Wzór na przedział ufności można wtedy zapisać w następujący sposób:

$$P\{\bar{x} - u_{\alpha} \frac{\sigma(x)}{\sqrt{n}} < \mu < \bar{x} + u_{\alpha} \frac{\sigma(x)}{\sqrt{n}}\} = 1 - \alpha$$

lub

$$P\{\bar{x} - u_{\alpha} \frac{s(x)}{\sqrt{n}} < \mu < \bar{x} + u_{\alpha} \frac{s(x)}{\sqrt{n}}\} = 1 - \alpha$$

gdzie: $\bar{x}$ - *średnia arytmetyczna z próby,*

u_{α}- *wielkość (parametr) odczytana z tablic dystrybuanty rozkładu normalnego,*

σ(x) – odchylenie standardowe z populacji,

s(x) - odchylenie standardowe z próby,

n – liczebność próby.

PRZEDZIAŁ UFNOŚCI DLA WARTOŚCI PRZECIĘTNEJ
MAŁA PRÓBA

Jeżeli liczebność próby jest n < 30, wówczas podstawą konstrukcji przedziału ufności jest statystyka: $t = \frac{\bar{x} - m}{s(x)} \sqrt{n-1}$ mająca rozkład t-Studenta o n-1 stopniach swobody.

Wzór na przedział ufności można wtedy zapisać w następujący sposób:

$$P\{\bar{x} - t_{\alpha,n-1} \frac{s(x)}{\sqrt{n-1}} < \mu < \bar{x} + t_{\alpha,n-1} \frac{s(x)}{\sqrt{n-1}}\} = 1 - \alpha$$

gdzie:

$t_{\alpha,n-1}$ - wartość odczytaną z tablic rozkładu t-Studenta przy odpowiednim α (poziomie istotności) i liczbie stopni swobody n – 1.

PRZEDZIAŁ UFNOŚCI DLA ODCHYLENIA STANDARDOWEGO

Jeżeli liczebność próby jest duża (n ≥ 30) wzór na przedział ufności można zapisać następująco:

$$P\{s(x)-u_{\alpha}\frac{s(x)}{\sqrt{2n}}<\sigma(x)<s(x)+u_{\alpha}\frac{s(x)}{\sqrt{2n}}\}=1-\alpha$$

Jeżeli liczebność próby jest mała (n < 30) estymację odchylenia standardowego przeprowadza się pośrednio, z wykorzystaniem przedziału ufności dla wariancji.

$$P\left\{\frac{\mathrm{ns}^2(\mathrm{x})}{\chi^2_{\frac{1}{2}\alpha,\mathrm{s=n-1}}}<\sigma^2(\mathrm{x})<\frac{\mathrm{ns}^2(\mathrm{x})}{\chi^2_{1-\frac{1}{2}\alpha,\mathrm{s=n-1}}}\right\}=1-\alpha$$

jeśli $S=\sqrt{\frac{1}{n}(x_i-\bar{x})^2}$

Czyli dla odchylenia standardowego otrzymujemy:

$$P\left\{\sqrt{\frac{\mathrm{ns}^2(\mathrm{x})}{\chi^2_{\frac{1}{2}\alpha,\mathrm{n-1}}}}<\sigma(\mathrm{x})<\sqrt{\frac{\mathrm{ns}^2(\mathrm{x})}{\chi^2_{1-\frac{1}{2}\alpha,\mathrm{n-1}}}}\right\}=1-\alpha$$

$\chi^2_{\frac{1}{2}\alpha,\mathrm{n-1}}$ - wielkości odczytane z tablic rozkladu chi - kwadrat

PRZEDZIAŁ UFNOŚCI DLA WSKAŹNIKA STRUKTURY

(PRAWDOPODOBIEŃSTWA SUKCESU, ODSETKA, FRAKCJI)

Przedział ufności dla frakcji buduje się na podstawie dużej próby ($n \geq 120$) wg wzoru:

$$P\left\{ w_i - u_\alpha \sqrt{\frac{w_i(1-w_i)}{n}} < p < w_i + u_\alpha \sqrt{\frac{w_i(1-w_i)}{n}} \right\} = 1-\alpha$$

gdzie: $w_i = \frac{n_i}{n}$, $\quad$ $\mathrm{n_i}$ – liczba wystąpień sukcesu w n - elementowej próbie.

w_i *można utożsamiać z prawdopodobieństwem sukcesu (częstość występowania sukcesu).*

MINIMALNA LICZEBNOŚĆ PRÓBY

W badaniach statystycznych często pojawia się problem określenia minimalnej liczebności próby potrzebnej do wyznaczenia określonego parametru z zadanym z góry maksymalnym błędem szacunku (d).

Problem ustalenia minimalnej wielkości próby może być rozstrzygany:

1. Przed badaniem, jako określenie minimalnej wielkości próby,
2. Po badaniu, gdy chodzi o odpowiedź na pytanie, czy ustalona liczebność próby jest wystarczająca, tzn. czy jest zgodna z poziomem żądanej precyzji wyników estymacji.

MINIMALNA LICZEBNOŚĆ PRÓBY

1. Jeśli szacujemy średnią μ ze zbiorowości generalnej (X: N(μ,σ) z błędem nie większym niż zadany (d) i znane jest σ w populacji generalnej, to wzór na minimalną liczebność próby przybiera postać:

$$n = \frac{u_\alpha^2 \sigma^2(x)}{d_{\bar{x}}^2}$$, $d_{\bar{x}}$ - maksymalny błąd szacunku

2. Dla nieznanego σ(x) i s(x) wyznaczonego na podstawie dużej próby pilotażowej można zapisać:

$$n = \frac{u_\alpha^2 s^2(x)}{d_{\bar{x}}^2}$$

3. Jeśli σ nie jest znane, a próba pilotażowa była mała, to wzór na minimalną liczebność próby przybierze postać:

$$n = \frac{t_{\alpha,n-1}^2 \cdot \hat{s}^2(x)}{d_{\bar{x}}^2}, \qquad \hat{s}^2(x) = \frac{1}{n-1}\sum_{1=1}^{n}(x_i - \bar{x})^2$$

MINIMALNA LICZEBNOŚĆ PRÓBY

4. Gdy szacujemy wskaźnik struktury niezbędną liczbę elementów n w próbie dla oszacowania p na poziomie ufności (1 – α) z maksymalnym błędem szacunku (d) wyznaczymy za pomocą wzoru:

a) jeśli rząd wielkości szacowanego prawdopodobieństwa *p* jest znany lub możemy go ocenić na podstawie wstępnej próby (n ≥ 120):

$$n = \frac{u_\alpha^2 p_i q_i}{d_{p_i}^2} = \frac{u_\alpha^2 w_i (1 - w_i)}{d_{p_i}^2}$$

b) jeśli nie znamy rzędu wielkości szacowanego parametru *p*, to wówczas zakładając wstępnie, że $p_i = q_i = 0{,}5$ otrzymujemy:

$$n = \frac{u_\alpha^2}{4d_{p_i}^2}$$

U!!! Jeśli obliczona liczebność próby jest ze względów praktycznych zbyt duża, to mniejszą liczebność otrzymamy zwiększając maksymalny błąd szacunku (zmniejszamy tym samym dokładność oszacowania).

U!!! Jeśli *n* nie jest liczbą całkowitą, to zaokrąglamy *n* w górę.

WERYFIKACJA (TESTOWANIE) HIPOTEZ STATYSTYCZNYCH

Testowanie hipotez statystycznych obejmuje zasady i metody określonych przypuszczeń (założeń) dotyczących parametrów lub postaci rozkładu cech statystycznych populacji generalnej na podstawie wyników z próby.

Hipotezą statystyczną nazywamy każdy sąd o zbiorowości generalnej (rozkładu cechy w populacji generalnej - hipotezy nieparametryczne lub wartości jego parametrów - hipotezy parametryczne) wydany bez przeprowadzenia badania całkowitego, jej prawdziwość orzeka się na podstawie próby losowej.

Inaczej mówiąc sądy (przypuszczenia) dotyczące populacji generalnej wydawane na podstawie próby nazywamy hipotezami statystycznymi. Wnioskowanie o słuszności prezentowanych sądów nazywamy sprawdzaniem lub weryfikacją hipotez statystycznych. Proces weryfikacji hipotezy przebiega według pewnego schematu postępowania nazywanego testem statystycznym. Testy na podstawie wyników z próby losowej pozwalają podjąć decyzję o odrzuceniu lub nie postawionej hipotezy.

Rodzaje hipotez statystycznych:

- *parametryczne* - dotyczą wartości parametrów populacji generalnej, takich jak wartość przeciętna, wariancja czy składnik struktury,
- *nieparametryczne* - dotyczą rozkładu zmiennej losowej w populacji generalnej (cechy statystycznej, współzależności cech, losowości próby).

WERYFIKACJA (TESTOWANIE) HIPOTEZ STATYSTYCZNYCH

Hipoteza zerowa (H_0) – hipoteza sprawdzana (weryfikowana) – formulowana często w taki sposób, aby na podstawie wyników z próby można było ją łatwo odrzucić (wbrew zdrowemu rozsądkowi), np. H_0: $\mu = \mu_0$ (hipoteza prosta) lub H_0: $\mu \geq \mu_0$, albo H_0: $\mu \leq \mu$ (hipotezy złożone).

Hipoteza alternatywna (H_1) – hipoteza, którą jesteśmy skłonni przyjąć, gdy odrzucimy hipotezę zerową.

Narzędziem umożliwiającym weryfikację hipotezy statystycznej jest tzw. *test statystyczny*.

Testem statystycznym nazywamy regułę postępowania określającą, przy jakich wynikach próby sprawdzaną hipotezę można przyjąć oraz przy jakich wynikach próby należy ją odrzucić.

Błąd I rodzaju (α) polega na odrzuceniu hipotezy H_0, mimo że jest ona prawdziwa.

Błąd II rodzaju (β) polega na przyjęciu H_0, gdy jest ona fałszywa.

W statystycznej kontroli jakości α określane jest często jako ryzyko producenta, β zaś jako ryzyko odbiorcy. Wartości α i β są wzajemnie powiązane, a zmniejszenie jednej powoduje zwiększenie drugiej.

Poziomem istotności α nazywamy prawdopodobieństwo popełnienia błędu I rodzaju. Najczęściej przyjmuje się wartości α równe: 0,01, 0,02, 0,05, 0,1.

WERYFIKACJA (TESTOWANIE) HIPOTEZ STATYSTYCZNYCH

Sprawdzian testu – zmienna losowa o określonym rozkładzie z próby, której wartość wpada lub nie do obszaru odrzucenia hipotezy zerowej (H_0), w zależności od tego, jaka będzie krytyczna wartość testu.

Dla hipotez parametrycznych sprawdzianami są estymatory odpowiednich parametrów, natomiast dla hipotez nieparametrycznych rolę sprawdzianów pełnią mierniki rozbieżności między rozkładem empirycznym a teoretycznym (sformułowanym w H_0).

Wartość krytyczna testu – wartość zmiennej losowej o określonym rozkładzie, która przy danym α (poziomie istotności) jest porównywana z wartością statystyki testu, dla ustalenia tego czy hipotezę zerową powinno się odrzucić czy przyjąć.

Zbiorem krytycznym nazywamy zbiór wartości sprawdzianu testu, które przemawiają za odrzuceniem H_0.

Rozkład sprawdzianu hipotezy określa, z jakich tablic należy odczytać wartość krytyczną wyznaczającą zbiór krytyczny. Zbiór krytyczny zależy zatem również od liczebności próby n, od tego, czy znamy parametry (μ lub σ) w zbiorowości generalnej oraz od poziomu istotności α.

WERYFIKACJA (TESTOWANIE) HIPOTEZ STATYSTYCZNYCH

Test jednostronny – sytuacja, w której zbiór krytyczny hipotezy zerowej znajduje się tylko na lewo lub tylko na prawo od wartości oczekiwanej danej zmiennej losowej. Zbiór krytyczny testu jest zatem usytuowany po jednej stronie wartości oczekiwanej.

Test dwustronny – sytuacja, w której zbiór krytyczny hipotezy zerowej umieszczony jest symetrycznie na lewo i na prawo od wartości oczekiwanej danej statystyki testu.

Moc testu – prawdopodobieństwo odrzucenia hipotezy zerowej, gdy hipoteza alternatywna jest prawdziwa. Moc testu zapisuje się następująco: $M = 1 – \beta$.

Zatem w procesie weryfikacji najlepszy byłby taki test, który zapewnia minimalizację α i β, a maksymalizuje moc testu. Taki test nosi nazwę *testu najmocniejszego*.

Poszukiwanie testów najmocniejszych nie jest łatwe, dlatego w praktyce badawczej wykorzystuje się testy, które uwzględniają tylko poziom istotności, czyli prawdopodobieństwo popełnienia błędu I rodzaju. Takie testy noszą nazwę *testów istotności*, gdy sprawdza się hipotezę parametryczną, bądź *testów zgodności*, gdy sprawdzana hipoteza jest nieparametryczna.

W konsekwencji uwzględnienie tylko poziomu istotności sprawia, że decyzja weryfikacyjna musi być ostrożniejsza. Ostrożność ta polega na tym, że na podstawie wyników próby nie podejmuje się decyzji w sposób kategoryczny typu: „przyjmujemy H_0” lub „odrzucamy H_0”, lecz raczej „odrzucamy H_0 na korzyść H_1, nie ma podstaw do odrzucenia H_0”.

WERYFIKACJA (TESTOWANIE) HIPOTEZ STATYSTYCZNYCH

Etapy weryfikacji hipotez statystycznych:

1. Ustalenie poziomu istotności
2. Postawienie H_0 i H_1 (w zależności od jej postawienia test może być jedno- lub dwustronny)
3. Wybór testu
4. Ustalenie sprawdzianu testu (statystyki) i jego wartości na podstawie dostępnych informacji o zbiorowości generalnej i próbie
5. Odczytanie wartości krytycznej sprawdzianu testu (najczęściej z tablic rozkładu normalnego, t-Studenta lub chi-kwadrat) oraz ustalenie obszaru krytycznego
6. Podjęcie decyzji weryfikacyjnej (na podstawie porównania wartości statystyki testu z wartością krytyczną) o odrzuceniu lub braku podstaw do odrzucenia hipotezy zerowej (*Jeżeli obliczona wartość testu znajduje się w obszarze krytycznym, to na poziomie istotności α H_0 należy odrzucić na korzyść hipotezy H_1. Można wtedy wnioskować, że wyniki badania empirycznego różnią się istotnie od założenia wyrażonego hipotezą zerową. Jeśli jednak obliczona wartość testu nie znajduje się w obszarze krytycznym, to nie ma podstaw do odrzucenia H_0. W tym przypadku można sądzić, że różnice między wynikami badania empirycznego, a założeniami danej hipotezy zerowej są statystycznie nieistotne*).

WERYFIKACJA HIPOTEZ PARAMETRYCZNYCH DOTYCZĄCYCH JEDNEJ ZBIOROWOŚCI GENERALNEJ

Test istotności dla nieznanej średniej (wartości przeciętnej) w populacji generalnej

Stawiamy odpowiednią hipotezę:

$H_0: \mu = \mu_0$

$H_1: \mu \neq \mu_0$ lub $H_1: \mu > \mu_0$ albo $H_1: \mu < \mu_0$

Zakładamy, ze rozkład cechy w zbiorowości generalnej jest $N(\mu, \sigma)$.

Wybór sprawdzianu hipotezy zależy od liczebności próby *n* oraz od tego, czy parametr σ w zbiorowości generalnej jest znany.

Jeśli znane jest $\sigma(x)$ *lub* gdy $n \geq 30$ (wtedy $\sigma = s$), to sprawdzianem H_0 jest statystyka:

$$u = \frac{\bar{x} - \mu_0}{\sigma(x)} \sqrt{n} \quad \text{lub} \quad u = \frac{\bar{x} - \mu_0}{s(x)}$$

gdzie: μ_0 - przypuszczalna średnia w populacji generalnej,

$s(x)$ - odchylenie standardowe z próby,

u - empiryczna wartość testu, którą porównuje się z wartością krytyczną u_α odczytaną z tablic dystrybuanty rozkładu normalnego na poziomie istotności α,

$\bar{x}$ - ocena μ_0 uzyskana z próby.

WERYFIKACJA HIPOTEZ PARAMETRYCZNYCH DOTYCZĄCYCH JEDNEJ ZBIOROWOŚCI GENERALNEJ

Test istotności dla nieznanej średniej (wartości przeciętnej) w populacji generalnej

Test istotności dla średniej *(wartości przeciętnej)* gdy $n < 30$

$$t = \frac{\bar{x} - \mu_0}{s(x)} \sqrt{n-1}$$

t - empiryczna wartość testu, którą należy porównać z wartością krytyczną $t_{\alpha,n-1}$ odczytaną z tablic rozkładu t-Studenta dla poziomu istotności α i n-1 stopni swobody.

!!! Jeśli relacja $t > t_\alpha$ jest spełniona, to H_0 należy odrzucić na korzyść H_1, w przeciwnym wypadku nie ma podstaw do odrzucenia H_0.

U!!!

W badaniu zawsze zakładamy, że populacja ma określony rozkład.

Najczęściej jest to rozkład normalny ze średnią μ i odchyleniem standardowym σ.

$$N(\mu, \sigma)$$

WERYFIKACJA HIPOTEZ PARAMETRYCZNYCH DOTYCZĄCYCH JEDNEJ ZBIOROWOŚCI GENERALNEJ

Test istotności dla wskaźnika struktury

Stawiamy odpowiednie hipotezy:

H_0: $p = p_0$

H_1: $p \neq p_0$ lub H_1: $p > p_0$ albo H_1: $p < p_0$

Sprawdzianem H_0 jest statystyka:

$$u = \frac{w_i - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

$$w_i = \frac{n_i}{n}$$

gdzie: n - liczebność próby,

p_0 - hipotetyczny wskaźnik struktury w populacji,

w_i - wskaźnik struktury w próbie,

u - empiryczna wartość testu, którą porównuje się z krytyczną wartością u_α

$q_0 = 1 - p_0$.

!!! Jeśli relacja $u > u_\alpha$ jest spełniona, to H_0 należy odrzucić na korzyść H_1, w przeciwnym wypadku nie ma podstaw do odrzucenia H_0.

WERYFIKACJA HIPOTEZ PARAMETRYCZNYCH DOTYCZĄCYCH DWÓCH ZBIOROWOŚCI GENERALNYCH

Test istotności dla dwóch średnich (wartości przeciętnych)

Chcemy zweryfikować hipotezę:

H_0: $\mu_1 = \mu_2$

wobec hipotezy:

H_1: $\mu_1 \neq \mu_2$ lub $\mu_1 > \mu_2$ albo $\mu_1 < \mu_2$

U!!! *Jeśli znane są odchylenia standardowe w populacjach σ_1 σ_2 i gdy próby liczą $n_1 < 30$, $n_2 < 30$ lub gdy odchylenia standardowe w populacjach σ_1 σ_2 nie są znane i próby liczą $n_1 \geq 30$, $n_2 \geq 30$, to do sprawdzianu hipotezy zerowej wykorzystujemy statystykę:*

$$u = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{\sigma_1^2(x)}{n_1} + \frac{\sigma_2^2(x)}{n_2}}} = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2(x)}{n_1} + \frac{s_2^2(x)}{n_2}}}$$

U!!! W niektórych źródłach uznaje się, że gdy $n_1 + n_2 > 122$ mamy do czynienia z dużą próbą.

gdzie : $\overline{x}_1, \overline{x}_2$ - średnie w próbie pierwszej i drugiej (1 i 2)

μ_1, μ_2 - hipotetyczne średnie w populacji 1 i 2

n_1, n_2 - liczebności próby 1 i 2

$\sigma_1^2(x), \sigma_2^2(x)$ - wariancje w populacji 1 i 2

$s_1^2(x), s_2^2(x)$ - wariancje w próbie 1 i 2

WERYFIKACJA HIPOTEZ PARAMETRYCZNYCH DOTYCZĄCYCH DWÓCH ZBIOROWOŚCI GENERALNYCH

Test istotności dla dwóch średnich (wartości przeciętnych)

U!!! Jeśli nieznane są odchylenia standardowe w populacjach i liczebność prób wynosi: $n_1 < 30$, $n_2 < 30$) , to do sprawdzianu hipotezy zerowej wykorzystujemy statystykę:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 s_1^2(x) + n_2 s_2^2(x)}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Statystyka ta (zmienna *t*) ma rozkład t-Studenta określony poziomem istotności α i liczbą stopni swobody $ss = n_1 + n_2 - 2$

!!! Jeśli relacja $t > t_\alpha$ jest spełniona, to H_0 należy odrzucić na korzyść H_1, w przeciwnym wypadku nie ma podstaw do odrzucenia H_0.

WERYFIKACJA HIPOTEZ PARAMETRYCZNYCH DOTYCZĄCYCH DWÓCH ZBIOROWOŚCI GENERALNYCH

Test istotności dla dwóch wariancji

Należy zweryfikować hipotezę:

$$H_0 : \sigma_1^2 = \sigma_2^2$$

przy $H_1 : \sigma_1^2 > \sigma_2^2$

Sprawdzianem hipotezy jest statystyka:

$$F = \frac{\hat{s}_1^2}{\hat{s}_2^2} \qquad \hat{s}^2(x) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

U!!! $\hat{s}_1 > \hat{s}_2$

Zmienna F ma rozkład F-Snedecora określony poziomem istotności α i liczbą stopni swobody $ss_1 = n_1 - 1$ i $ss_2 = n_2 - 1$.

!!! Jeśli relacja $F > F_\alpha$ jest spełniona, to H_0 należy odrzucić na korzyść H_1, w przeciwnym wypadku nie ma podstaw do odrzucenia H_0.

F_α odczytane z tablic obliczane są dla prawostronnego obszaru krytycznego, dlatego nie należy podwajać poziomu istotności.

Przedstawiony powyżej test wykorzystuje się najczęściej do sprawdzenia prawdziwości założenia: $\sigma_1 = \sigma_2$, które musi być spełnione (jest konieczne), aby można było testować hipotezę o równości wartości przeciętnych w dwóch populacjach, gdy $n_1 < 30$, $n_2 < 30$, a σ_1 i σ_2 są nieznane.

!!! Jeśli w zadaniach dane są wartości wariancji s^2, to wartości, $\hat{s}^2$ z których korzystamy w teście wyznacza się z równości:

$$\hat{s}^2 = \frac{ns^2(x)}{n-1}$$

WERYFIKACJA HIPOTEZ PARAMETRYCZNYCH DOTYCZĄCYCH DWÓCH ZBIOROWOŚCI GENERALNYCH

Test istotności dla dwóch wskaźników struktury

Należy zweryfikować hipotezę:

$H_0: p_1 = p_2$

przy $H_1: p_1 \neq p_2$ lub $H_1: p_1 > p_2$ albo $H_1: p_1 < p_2$

Sprawdzianem hipotezy H_0 jest statystyka:

$$u = \frac{w_1 - w_2}{\sqrt{\frac{\overline{p}\overline{q}}{\overline{n}}}}$$

$$w_i = \frac{m_i}{n_i}$$

gdzie: w_1, w_2 - wskaźniki struktury w próbie 1 i 2,

p_1, p_2 - wskaźniki struktury w populacji 1 i 2

$$\overline{p} = \frac{m_1 + m_2}{n_1 + n_2}$$

m_1 - liczba wyróżnionych elementów w próbie 1

m_2 - liczba wyróżnionych elementów w próbie 2

$$\overline{n} = \frac{n_1 \cdot n_2}{n_1 + n_2}$$

n_1 - liczebność próby 1

n_2 - liczebność próby 2

Obydwie próby musza być duże, tzn. $n_1 > 100$ i $n_2 > 100$.

WERYFIKACJA HIPOTEZ NIEPARAMETRYCZNYCH
WYBRANE TESTY

Nieparametryczne testy istotności dzielimy na trzy zasadnicze grupy:

• testy zgodności,

• testy niezależności,

• testy losowości próby.

Testy nieparametryczne, w przeciwieństwie do testów parametrycznych, mają tę zaletę, że nie wymagają założeń w odniesieniu do postaci rozkładu cechy w zbiorowości generalnej.

Test zgodności χ^2

Test zgodności chi-kwadrat służy do weryfikacji hipotezy, że zaobserwowana cecha X w zbiorowości generalnej ma określony typ rozkładu np. dwumianowy, Poissona, normalny.

Narzędziem weryfikującym hipotezę, że cecha ma określony typ rozkładu jest test zgodności χ 2. Hipotezę zerową można formułować wykorzystując następujące sposoby:

H_0: cecha X ma rozkład określony dystrybuantą $F(x) = F_0(x)$,

H_0: cecha X ma rozkład $N(\mu, \sigma)$,

H_0: cecha X ma rozkład $N(10, 1)$,

H_0: cecha X ma rozkład Poissona itp.

Stawiamy odpowiednie hipotezy:

$$H_0 : F(x) = F_0(x)$$

$$H_1 : F(x) \neq F_0(x)$$

$F(x)$ - dystrybuanta rozkładu populacji,

$F_0(x)$ - dystrybuanta zakładanego typu rozkładu

WERYFIKACJA HIPOTEZ NIEPARAMETRYCZNYCH
WYBRANE TESTY

Test zgodności χ^2

Sprawdzianem hipotezy H_0 jest statystyka:

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - \hat{n}_i)^2}{\hat{n}_i}$$

gdzie: n_i - liczebności empiryczne w rozkładzie empirycznym dla i-tego wariantu cechy,

$\hat{n}_i$ - liczebnośi i teoretyczne w rozkłozkła teoretycznym dla i - tego wariantu zmiennej,

które oblicza siê korzystaj¹c ze wzoru o postaci : $\hat{n}_i = n \cdot p_i$

Zmienna χ^2 w przypadku prawdziwości H_0 ma rozkład χ^2 określony liczbą stopni swobody ss = k - r - 1

gdzie: k - liczba przedziałów klasowych (wymóg co najmniej 5 elementów w każdym z przedziałów),

r - liczba parametrów, które należy wstępnie wyznaczyć na podstawie próby.

!!! Jeśli relacja $\chi > \chi_\alpha$ jest spełniona, to H_0 należy odrzucić na korzyść H_1, w przeciwnym wypadku nie ma podstaw do odrzucenia H_0.

$$\hat{n}_i = n \cdot p_i$$

We wzorze p_i oznacza, że cecha X przyjmuje wartość należącą do i-tego przedziału klasowego (gdy rozkład cechy jest zgodny z H_0), np_i zaś oznacza liczbę jednostek, które powinny znaleźć się w i-tym przedziale, przy założeniu, że cecha ma rozkład zgodny z hipotetycznym (określa się je jako liczebności teoretyczne). Statystyka χ^2 jest miarą rozbieżności między rozkładem empirycznym a teoretycznym, (duże wartości χ^2 powodują odrzucenie H_0.)

WERYFIKACJA HIPOTEZ NIEPARAMETRYCZNYCH
WYBRANE TESTY

Test losowości próby - *test serii*

Serią nazywamy każdy podciąg złożony z kolejnych elementów jednego rodzaju, utworzony w ciąg uporządkowanych w dowolny sposób elementów dwojakiego rodzaju.

Stawiamy odpowiednie hipotezy: H_0: próba jest losowa,

H_1: próba nie jest losowa

Procedura testowa opiera się na następujących krokach:

1. Wyniki losowania porządkujemy by wyznaczyć medianę.
2. Każdemu wynikowi losowania uwzględniając chronologię wyboru przypisuje się symbol „a”, jeżeli $x_i < M_e$, bądź symbol „b”, jeżeli $x_i > Me$. Wynik $x_i = Me$ pomija się.
3. Dla ciągu symboli a i b podaje się liczbę serii k, która ma znany i stablicowany rozkład zależny od n_1 i n_2 i liczebności elementów a oraz b.
4. Z tablic rozkładu serii dla poziomu istotności α oraz liczebności n_1 i n_2 odczytuje się dwie wartości krytyczne: k_1, dla której $P(k \leq k_1) = ½\, \alpha$ i k_2, dla której $P(k \leq k_2) = 1 - ½\, \alpha$.
5. Podejmowanie decyzji: jeżeli $k \leq k_1$ lub $k \geq k_2$, to H_0 odrzuca się. Natomiast jeżeli $k_1 < k < k_2$, nie ma podstaw do odrzucenia H_0, co oznacza, że próba jest losowa.

Test losowości próby - *test serii*

Przykład:

Maszyna produkuje detale o określonej średnicy. Do kontroli dokładności wymiarów wylosowano kolejno 16 sztuk, otrzymując następujące wyniki pomiarów (w mm):
8,8 9,2 10,1 10,0 9,7 10,6 11,8 10,5 12,1 11,5 9,9 12,6 12,4 12,8 13,0 12,7 . Na poziomie istotności 0,05 sprawdzić hipotezę, że wybór dał próbę losową.

Rozwiązanie:

Me dla uporządkowanych danych wynosi:

(10,6 + 11,5):2 = 11,05

Ciąg symboli a i b jest następujący: aaaaaababbabbbbb,

stąd k = 6. Z tablic rozkładu serii przy $n_1 = 8$ i $n_2 = 8$ odczytujemy:

$k_1 = 4$, a $k_2 =13$.

Ponieważ $k_1 = 4 < k=6 < k_2 = 13$,

… nie ma podstaw do odrzucenia H_0, co oznacza, że na poziomie istotności 0,05 *wylosowana próba jest losowa*.