Глава 13

Проверка на статистически хипотези

13.1 Понятие за статистическа хипотеза

В много научни изследвания задачата може да се формулира във вида на хипотеза, която предстои да бъде потвърдена или отхвърлена. В статистическата практика се налага да се проверяват твърдения относно стойностите на генералните параметри или закона за разпределение на една или няколко статистически променливи. Такива твърдения се наричат **статистически хипотези**. Например, статистически са следните хипотези:

1. Нормално разпределена статистическа променлива има дадени средно и дисперсия;

2. Нормално разпределение има дадено средно (за дисперсията не се твърди нищо);

3. Разпределението на статистическа променлива е нормално (с някакви средно и дисперсия);

4. Две статистически променливи имат еднакво разпределение (за вида на закона за разпределение не се твърди нищо).

В случая на една статистическа променлива се проверяват два основни типа статистически хипотези:

• проверка на хипотезата за равенство на стойността на даден генерален параметър на определено хипотетично число;

• проверка на хипотезата за разпределение на изследваната статистическа променлива по даден вероятностен закон.

13.2 Видове статистически хипотези

13.2.1 Параметрични и непараметрични хипотези

Ще отбележим,че в примерите 1 и 2 на предната точка разпределението има известен вид (то е нормално) и хипотезите са само за стойността на единия или двата параметъра на това разпределение. Такива хипотези се наричат **параметрични**. Хипотезите от примерите 3 и 4 имат друга природа. Те са относно вида на закона за разпределение на една или няколко статистически променливи и не са свързани със стойностите на конкретни параметри на тези разпределения. Такива хипотези се наричат **непараметрични**.

13.2.2 Прости и сложни хипотези

Между хипотезите от примерите 1 и 2 също има разлика. Хипотезата от пример 1 е относно стойностите на всичките параметри на разпределението (в случая два параметъра), докато хипотезата от пример 2 е относно само един от тези два параметъра. За теорията това различие е съществено. В общия случай то се формулира по следния начин. Ако разпределението има общо r параметъра и хипотезата твърди, че k от тях имат определени стойности, тя се нарича проста, ако $k = r$, и сложна, ако $k < r$. При това числото $r - k$ се нарича **брой на степени на свобода** на хипотезата, а числото k – **брой на ограниченията**, налагани от хипотезата.

13.2.3 Основна и алтернативна хипотеза

Проверяваната хипотеза се нарича **основна (работна, нулева)** и се бележи с H_0. Наред с хипотезата H_0 се разглежда и хипотеза H_1, която ѝ противоречи и се нарича **алтернативна (конкурираща)**. Например, ако се разглежда хипотеза за равенство на параметър θ на някакво хипотетично число θ_0, т.е. $H_0 : \theta = \theta_0$, то като алтернативни може да разгледаме една от следващите хипотези:

$$H_1^{(1)} : \theta > \theta_0, \quad H_1^{(2)} : \theta < \theta_0, \quad H_1^{(3)} : \theta \neq \theta_0,$$

$$H_1^{(4)} : \theta = \theta_1 \ \ (\theta_1 \text{ е дадена стойност и } \theta_1 \neq \theta_0).$$

Изборът на алтернативната хипотеза се определя от конкретната формулировка на задачата за проверка на основната хипотеза и решението за приемане или отхвърляне на основната хипотеза съществено зависи от това, против каква алтернативна хипотеза тя се проверява.

13.3 Статистическа проверка на хипотези. Грешки от първи и втори род

Основната хипотеза може да бъде вярна или грешна и затова възниква необходимостта от нейната проверка. Статистическа проверка на хипотеза се нарича процедурата по изясняване дали да се приеме основната хипотеза или да се отхвърли, при което решението за приемане или отхвърляне се базира на данните от случайна статистическа извадка:

$$x_1, x_2, \ldots, x_n. \tag{13.1}$$

В хода на статистическата проверка на произволна статистическа хипотеза с критерий за проверка, се открояват следните основни етапи:

1. Определяне на основната (H_0) и алтернативната (H_1) хипотези на проверката.

2. Определяне на нивото на значимост α на проверката.

3. Избиране на критерий за проверка на основна хипотеза.

4. Определяне на критичната област W на критерия за проверка.
   или
   Пресмятане на $P-$стойността на критерия за проверка.

5. Вземане на решение относно статистическата основателност на основната хипотеза.

Първо ще се спрем по-подробно на основните етапи на проверката в случая, когато се използва критична област, а след това ще отбележим с какво се отличават етапите 4 и 5 в случая, когато се използва $P-$стойност.

Причината за обособяването на основната хипотеза се състои в това, че H_0 обикновено се разглежда като твърдение, което е по-важно, ако бъде отхвърлено. Това се основава на общия принцип, че теорията трябва да бъде отхвърлена, ако има поне един противоречащ ѝ пример, но не е задължително да бъде приета, ако такъв пример не е намерен. Понеже решението за приемане или отхвърляне на H_0 се основава на данните на случайната извадка (13.1), то това решение може да се окаже грешно. Възможно е да се допуснат два рода грешки:

Ако хипотезата H_0 е вярна, а е прието решение за нейното отхвърляне, то казваме, че се допуска **грешка от първи род**.

Ако хипотезата H_1 е вярна, а е прието решение за приемане на H_0, то казваме, че се допуска **грешка от втори род**.

Вероятността за грешка от първи род се означава с α, а вероятността за грешка от втори род се означава с β. Тези вероятности може да се представят във вида

$$\alpha = P(\text{да се отхвърли } H_0 | \ H_0 \text{ е вярна}), \tag{13.2}$$

$$\beta = P(\text{да се приеме } H_0 | \ H_1 \text{ е вярна}). \tag{13.3}$$

Ще подчертаем, че последствията от тези грешки може да се отличават съществено. Може да се посочат примери, когато грешката от първи род влече по-тежки последствия, отколкото грешката от втори род.

При проверка на хипотеза може да се приеме правилно решение в следните два случая:

1) Ако хипотезата H_0 е вярна и е прието решение за приемане на H_0.
2) Ако хипотезата H_1 е вярна и е прието решение за отхвърляне на H_0.

Казаното дотук относно различните възможни изходи от статистическата проверка на статистическите хипотези, може да се представи нагледно с помощта на следната таблица:

Два типа грешки, допускани при статистическа проверка на хипотеза

	H_0 е вярна	H_1 е вярна
H_0 е отхвърлена	Грешка от първи род, вероятност α	Вярно решение, вероятност $\gamma = 1 - \beta$
H_0 е приета	Вярно решение, вероятност $p = 1 - \alpha$	Грешка от втори род, вероятност β

Величината $p = 1 - \alpha$ се нарича **доверителна вероятност** на проверката и показва каква е вероятността да приемем основната хипотеза, в случай че тя е вярна. Величината $\gamma = 1 - \beta$ се нарича **мощност на проверката** и показва каква е вероятността да отхвърлим основната хипотеза H_0 при условие, че е вярна алтернативната хипотеза H_1. Не е трудно да се провери, че

$$\gamma = P(\text{да се отхвърли } H_0 |\ H_1 \text{ е вярна}). \qquad (13.4)$$

13.4 Статистически критерий за проверка. Наблюдавана стойност на критерия

Правилото $\mathcal{K}$, по което се взема решение за приемане или отхвърляне на основната хипотеза H_0, се нарича **статистически критерий** за проверка (или просто **критерий**). За да се формулира конкретен критерий за проверка, се използва специално подбрана функция на данните от наблюдението (извадката (13.1)), например $Z = Z(X_1, X_2, \ldots, X_n)$. Тази функция се нарича **статистика Z на критерия $\mathcal{K}$**.

Тъй като данните от извадката са статистически реализации на случайна величина X с определен закон за разпределение, то статистиката Z също е случайна величина с определен закон за разпределение. За статистика на критерия за проверка на една статистическа хипотеза винаги се избира такава случайна величина $Z = Z(X_1, X_2, \ldots, X_n)$, чието разпределение (функция на разпределение или плътност на разпределение) е известно (точно или приближено), **при предположение, че проверяваната основна хипотеза H_0 е вярна**. Тази величина се означава с U, ако е разпределена нормално; с F, ако е разпределена по закона на Фишер–Снедекор; с T, ако е разпределена по закона на Стюдънт; с χ^2, ако е разпределена по закона „хи-квадрат” и т.н. Понеже в този параграф няма да конкретизираме вида на избраната статистика за проверка, ще я означаваме общо с Z.

За всеки конкретен статистически метод за проверка на основната хипотеза, който е реализиран програмно, изборът на статистиката на критерия вече е направен.

Наблюдавана стойност Z_0 на критерия се нарича стойността на статистиката Z, която е изчислена по извадката (13.1):

$$Z_0 = Z(x_1, x_2, \ldots, x_n).$$

Например, ако се проверява хипотезата за равенство на дисперсиите на две нормално разпределени генерални съвкупности, то за статистика на критерия се приема отношението $F = s_1^2/s_2^2$ на неизместените статистически дисперсии. Тази величина е случайна, понеже се изчислява по стойностите на различните наблюдения (извадки) и нейното разпределение е по закона на Фишер–Снедекор. За проверка на хипотезата по данните на извадките първо се изчисляват частните стойности на величините s_1^2 и s_2^2, участващи в статистиката, след което се намира наблюдавана стойност на критерия. Например, ако по две извадки са намерени неизместените статистически дисперсии $s_1^2 = 24$ и $s_2^2 = 16$, то наблюдаваната стойност на критерия е $F_0 = s_1^2/s_2^2 = 24/16 = 1.5$.

13.5 Критична област. Област на приемане на хипотезата. Критични точки

След избора на статистиката Z на критерия, множеството от всички възможни стойности на Z се разбива на сума на две непресичащи се множества W и $\overline{W}$. Множеството W се нарича **критична област** и се състои от всички стойности на Z, при които основната хипотеза се отхвърля. Множеството $\overline{W}$ се нарича **област на допустими стойности** и се състои от всички стойности на Z, при които основната хипотеза се приема.

Когато статистическата проверката на основната хипотеза се извършва с използване на критична област, решението за приемане или отхвърляне на тази хипотеза се взема съгласно следното правило.

Правило за проверка: Ако наблюдаваната стойност на критерия Z_0 попада в критичната област W, хипотезата H_0 се отхвърля; ако Z_0 попада в допустимата област $\overline{W}$, то хипотезата H_0 се приема.

Забележка 13.1. *Нека основната хипотеза H_0 е приета; грешно е да се мисли, че с това тя е доказана. Наистина, известно е, че един пример, потвърждаващ справедливостта на някакво общо твърдение, все още не го доказва. Затова е по-правилно да се казва, че „данните от наблюдението се съгласуват с основната хипотеза и не дават основание за нейното отхвърляне". На практика, за да приемем основната хипотеза с по-голяма увереност, същата се проверява по друг начин или проверката се повтаря с по-голям обем на извадката.*

Отхвърлянето на основната хипотеза става по-категорично, отколкото нейното приемане. Наистина, достатъчно е да се посочи само един пример, който противоречи на някакво общо твърдение, за да бъде отхвърлено това твърдение. Ако се е оказало, че наблюдаваната стойност на критерия принадлежи на критичната област, то този факт служи за пример, противоречащ на основната хипотеза, което е основание за нейното отхвърляне.

При определянето на критичната област W първо се задава положително число α, което се нарича **ниво на значимост**. Нивото на значимост е най-голямата вероятност, с която приемаме при проверката да се допусне грешка от първи род. Обикновено α се избира от провеждащия проверката и приема малки стойности: 0.05, 0.01, 0.001. Например, ако изберем ниво на значимост $\alpha = 0.05$, то това означава, че приемаме средно при 5 от 100 проверки на дадена хипотеза да се допусне грешка от първи род (да отхвърлим правилна хипотеза).

След избора на α се търси такава критичната област W, при която величината

$$\gamma = P(Z_0 \text{ попада в } W \mid H_1 \text{ е вярна}) \text{ е максимална} \tag{13.5}$$

при условие, че

$$P(Z_0 \text{ попада в } W \mid H_0 \text{ е вярна}) \le \alpha\,. \tag{13.6}$$

Условие (13.6) означава, че статистиката Z трябва да бъде подбрана така, че вероятността за допускане на грешка от първи род да не превишава избраното ниво на значимост α.

Условие (13.5) означава, че статистиката Z трябва да бъде подбрана така, че мощността на проверката γ да е максимална (при ограничението (13.6)), т.е. при това ограничение вероятността β за допускане на грешка от втори род да е минимална.

Забележка 13.2. *Ясно е, че колкото вероятностите за грешки от първи и втори род са по-малки, толкова критичната област е „по-хубава". Обаче, при фиксиран обем на извадката е невъзможно да се намалят едновременно α и β; ако намаляваме α, то β ще нараства. Възниква въпросът как най-целесъобразно да се избере α? Отговорът на този въпрос зависи от „тежестта на последствията" от двата рода грешки във всяка конкретна задача. Например, ако грешката от първи род* *ще доведе до по-големи загуби, а от втори род – до по-малки, то трябва да се приеме възможно по-малко* α*.*

Забележка 13.3. *Единственият способ за едновременно намаляване на вероятностите за грешки от първи и втори род е чрез увеличаване на обема на извадката, който обаче не може да расте безгранично.*

Понеже статистиката Z на критерия $\mathcal{K}$ е едномерна случайна величина, то нейните всевъзможни стойности принадлежат на някакъв интервал. Затова областите W и $\overline{W}$ също са интервали и следователно съществуват точки, които ги разделят.

Критични точки (граници) c_{kp} наричаме точките, отделящи критичната област W и областта на допустимите стойности $\overline{W}$.

Критичните области биват:

- **едностранна дясна критична област** $[c_r, +\infty)$, която се задава с неравенството $Z \geq c_r > 0$;

- **едностранна лява критична област** $(-\infty, c_l]$, която се задава с неравенството $Z \leq c_l < 0$;

- **двустранна критична област** $(-\infty, c_1] \cup [c_2, +\infty)$, която се задава с неравенствата $Z \leq c_1$ и $Z \geq c_2 \quad (c_1 < c_2)$.
  В частност, ако двустранната критична област е симетрична относно нулата $(-c_1 = c_2)$, то тя се задава с неравенство от вида $|Z| \geq c \quad (c > 0)$.

Ще отбележим, че от алтернативната хипотеза H_1 зависи какъв да бъде видът на избраната критична област. Например, ако се разглежда хипотеза за равенство на параметър θ на някакво хипотетично число θ_0, т.е. $H_0 : \theta = \theta_0$, то при алтернативна хипотеза $H_1 : \theta > \theta_0$ се избира едностранна дясна критична област, при алтернативна хипотеза $H_1 : \theta < \theta_0$ се избира едностранна лява критична област, а при алтернативна хипотеза $H_1 : \theta \neq \theta_0$ се избира двустранна критична област.

13.6 Намиране на критичната област. Вземане на решение

След като е определен видът на критичната област, нейното намиране се извършва, като се реши оптимизационната задача (13.5)–(13.6). Начините за нейното решаване не са лесни и затова няма да се спираме върху тях. Ще отбележим, че за най-често срещаните задачи за проверка на хипотези е направен подходящ избор на статистика на критерия и са определени критичните области, като са намерени критичните точки. За всички конкретни статистически методи за проверка, които са реализирани програмно, това вече е направено.

Ще отбележим също, че намирането на критичните точки е свързано с решаване на следните уравнения:

$$P(Z \geq c_r | \ H_0 \text{ е вярна}) = \alpha. \tag{13.7}$$

(за едностранна дясна критична област);

$$P(Z \leq c_l | \ H_0 \text{ е вярна}) = \alpha. \tag{13.8}$$

(за едностранна лява критична област);

$$P(Z \leq c_1 | \ H_0 \text{ е вярна}) = \frac{\alpha}{2} \tag{13.9}$$

и

$$P(Z \geq c_2 | \ H_0 \text{ е вярна}) = \frac{\alpha}{2} \cdot \tag{13.10}$$

(за двустранна критична област).

За да се вземе решение за приемане или отхвърляне на основната хипотеза, се изчислява наблюдаваната стойност $Z_0 = Z(x_1, x_2, \ldots, x_n)$ и съгласно правилото за проверка се вземат следните решения:

– за едностранна дясна критична област:
При $Z_0 \geq c_r$, хипотезата H_0 се отхвърля, а при $Z_0 < c_r$ хипотезата H_0 се приема;

– за едностранна лява критична област:
При $Z_0 \leq c_l$, хипотезата H_0 се отхвърля, а при $Z_0 > c_l$ хипотезата H_0 се приема;

– за двустранна критична област:
При $Z_0 \leq c_1$ или $Z_0 \geq c_2$ хипотезата H_0 се отхвърля, а при $c_1 < Z_0 < c_2$ хипотезата H_0 се приема.

13.7 Понятие за $P-$стойност. Вземане на решение

Разгледаната по-горе процедурата за вземане на решение за приемане или отхвърляне на основната хипотеза се осъществява с построяване на критичната област на критерия. В практиката се използва и еквивалентна на нея процедура, която се осъществява с пресмятането на т.нар. $P-$стойност на критерия. Нека при дадени основна (H_0) и алтернативна (H_1) хипотези и избрано ниво на значимост α е подбрана статистика Z на критерия за проверка $\mathcal{K}$ и е определен вида на критичната област. Нека по данните на извадката (13.1) е пресметната наблюдаваната стойност $Z_0 = Z(x_1, x_2, \ldots, x_n)$ на критерия. $P-$стойността на критерия $\mathcal{K}$ е вероятност, която се определя по една от формулите:

– за едностранна дясна критична област:

$$P = P(Z \geq Z_0 |\ H_0 \text{ е вярна}); \tag{13.11}$$

– за едностранна лява критична област:

$$P = P(Z \leq Z_0 |\ H_0 \text{ е вярна}); \tag{13.12}$$

– за двустранна критична област:

$$P = 2\min(P(Z \leq Z_0 |\ H_0 \text{ е вярна}), P(Z \geq Z_0 |\ H_0 \text{ е вярна}))\,, \tag{13.13}$$

където $\min(a, b)$ означава по-малкото от числата a, b.

За изхода от проверката са възможни два различни случая:

1) За $P-$стойността на критерия за проверка е в сила: $P \leq \alpha$.

Тогава е налице статистически значим резултат. Основната хипотеза H_0 се отхвърля, защото са налице основателни статистически доводи за това;

2) За $P-$стойността на критерия за проверка е в сила: $P > \alpha$.

Тогава е налице статистически незначим резултат. Основната хипотеза H_0 се приема, защото липсват основателни статистически доводи за нейното отхвърляне.

Вземането на решение за приемане или отхвърляне на основната хипотеза чрез пресмятане на $P-$стойността на критерия се налага в практиката напоследък, поради възможността за бързо и точно пресмятане на тази стойност с помощта на компютър.