Бизнис статистика

Тестирање на хипотези – едно обележје

Вовед

- Точкастото оценување и оценувањето со интервали на доверба се важни компоненти од статистичката анализа.
- Но, во многу проблеми во пракса треба да се одлучи дали се прифаќа или отфрла некое тврдење во врска со некој параметар или распределба.
- Процедурата за донесување на одлука дали една претпоставка е точна или не се нарекува ***тестирање на хипотези.***
- Ова е една од најкористените алатки за статистичко заклучување.
- Подоцна, ќе видиме дека постои тесна врска помеѓу тестирање на хипотези и интервали на доверба.

Примери кои вклучуваат тестирање на хипотези

- Сакаме да провериме дали една паричка е фер, т.е. дали $p = 0.5$ кај биномната распределба?
  - За таа цел, паричката се фрла n пати и се формира примерок од добиените исходи. Врз основа на добиените резултати, треба да се спроведе статистички тест во кој треба да се одлучи дали претпоставката (хипотезата) дека паричката е фер ќе се прифати или ќе се отфрли.
- Просечното време на чекање на клинетите во една банка било 5 минути. Сакаме да провериме дали дошло до намалување на просечното време на чекање откако се вовел нов шалтер?
  - За таа цел, се избираат случајно одреден број на клиенти и се формира промерок од нивните времиња на чекање. Врз основа на добиените резултати, треба да се спроведе статистички тест во кој треба да се одлучи дали претпоставката (хипотезата) дека времето на чекање е намалено ќе се прифати или ќе се отфрли.

Статистички хипотези

Дефиниција 1. Секоја претпоставка за популацијата, за распределбата на некое нејзино обележје, за вредностите на некои параметри на обележјето и сл. се вика ***статистичка хипотеза***.

- Постапката со која врз основа на податоците од даден примерок се проверува точноста на поставената хипотеза се нарекува ***статистички тест***.

Параметарски и непараметарски хипотези

- Доколку е позната распределбата на обележјето, но непознати се параметрите од кои зависи таа распределба и во хипотезата се претпоставени некои вредности за непознатите параметри, таквата хипотезата се нарекува ***параметарска***. Во секој друг случај велиме дека хипотезата е ***непараметарска***.
- Еве неколку примери:
  - Нека е дадено дека случајна променлива X има нормална распределба. Тогаш хипотезата "Математичкото очекување на X е 0" е параметарска хипотеза и се однесува на μ во $N(\mu,\sigma^2)$.
  - Нека е дадено дека X и Y имаат нормална распределба. Тогаш хипотезата "Дисперзиите на X и Y се еднакви" е параметарска хипотеза.
  - "Случајната променлива X има нормална распределба" е непараметарска хипотеза.
  - "Случајната променлива X има Пуасонова распределба $P(1)$" е непараметарска хипотеза.
  - "Случајните променливи X и Y се независни" е непараметарска хипотеза.

Статистички хипотези

- Статистички тест е всушност постапка која се сведува на одлучување за една од две хипотези кои се исклучуваат една со друга. Ако едната се прифати, другата автоматски се отфрла, и обратно.

- Така, во примерот со фер паричка, едната хипотеза е дека паричката е фер или веројатноста да се појави глава (а со тоа и петка) е ½, а втората хипотеза е дека таа веројатност е различна од ½.

- Во вториот пример со време на чекање во банка, едната хипотеза е дека времето на чекање ќе остане 5 минути, а втората е дека времето на чекање ќе се намали (со воведување на новиот шалтер).

Статистички хипотези

- Хипотезата која се тестира се нарекува ***нулта хипотеза*** и се означува со H_0, а спротивната се нарекува ***алтернативна хипотеза*** и се означува со H_a.
  - Нултата и алтернативната хипотеза се дисјунктни.
- За нулта хипотеза обично се избира онаа во која се зададени некои стандарди за конкретната карактеристика или вредностите кои биле актуелни до последното мерење.
  - Така, за првиот пример, ако p е веројатноста за појава на глава, тогаш

$$H_0: p = ½$$

$$H_a: p \neq ½$$

  - За вториот пример, ако μ_1 е очекувано време на чекање пред воведување на нов шалтер, а μ_2 е очекувано време на чекање по воведување на нов шалтер, тогаш

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 > \mu_2$$

Параметарски хипотези

- Во понатамошниот дел од ова предавање ќе се задржиме на тестирање на параметарски хипотези.
- Примерите кои се дадени на претходниот слајд се примери за параметарски хипотези.
- Да воочиме дека статистичките хипотези се однесуваат на параметарот на распределбата на обележјето, а не на статистиката на примерокот.

$$H_0 : \mu \geq 3$$

$$\cancel{H_0 : \bar{x} \geq 3}$$

Параметарски хипотези

- Хипотезите во кои се појавува само една вредност на непознатиот параметар се нарекуваат ***прости хипотези***.
  - Така, хипотезата H_0: $p = ½$ е проста хипотеза.

- Хипотезата со која за параметарот се допуштени повеќе вредности на параметарот се нарекува ***сложена хипотеза***.
  - Хипотезата H_a: $p \neq ½$ е сложена хипотеза, затоа што тука за p се допуштени сите вредности од интервалот (0,1) кои се различни од ½.

Еднострaн и двостран тест

- Во примерот за фер паричка, ги поставивме следните хипотези:

$$H_0: p = ½$$
$$H_a: p \neq ½$$

- Бидејќи алтернативната хипотеза е H_a: $p \neq ½$, ваквиот тест се нарекува ***двостран тест***.

- Во некои случаи, хипотезите може да се поставаат како

$$H_0: p = ½ \qquad\qquad H_0: p = ½$$
$$H_a: p > ½ \qquad \text{или} \qquad H_a: p < ½$$

- Во тој случај, велиме дека тестот е ***еднострaн***.

Тестирање на хипотези

- Процедурата што доведува до одлучување за одредена хипотеза се нарекува ***тестирање на хипотези***.
- Постапките за тестирање на хипотезите се потпираат на користење на информациите кои може да се добијат од еден случаен примерок земен од популацијата од интерес.
- Ако овие информации се во согласност со хипотезата, ќе заклучиме дека хипотезата е ***точна***. Обратно, ако овие информации не се во согласност со хипотезата, ќе заклучиме дека хипотезата ***не е точна***.
- Треба да се нагласи дека точноста на една хипотеза никогаш не може да се знае со сигурност, освен ако не можеме да ја испитаме целата популација, што е невозможно во повеќе практични ситуации.
- Затоа, процедурите за тестирање на хипотези се развиваат со одредена веројатност да се донесе погрешен заклучок.

Грешка од прв и втор тип

- Погрешен заклучок би се донел, ако ја отфрлиме нултата хипотеза кога е точна или ако ја прифатиме нултата хипотеза кога е таа е неточна (т.е. точна е алтернативната хипотеза). Затоа дефинираме два типа на грешки.

Дефиниција 2. Грешката што се прави ако се отфрли нултата хипотеза кога таа е точна се вика ***грешка од прв тип.*** Грешката што се прави ако се прифати нултата хипотеза кога таа не е точна се нарекува ***грешка од втор тип***.

Грешка од I тип и грешка од II тип

- Според тоа при било кое тестирање на статистички хипотези можни се четири различни ситуации кои одредуваат дали конечното решение е точно или погрешно. Истите се дадени во табелата

	H_0 **е точна**	H_0 **не е точна**
Тестот ја отфрла H_0	**Грешка од I тип**	Нема грешка
Тестот не ја отфрла H_0	Нема грешка	**Грешка од II тип**

Грешка од I тип и грешка од II тип

- Бидејќи нашето решение е врз основа на случајни променливи, на грешките од тип I и тип II можеме да им придружиме веројатности. Овие веројатности ги означуваме со α и β соодветно, односно

$$\alpha = P(\text{грешка од тип I}) = P(H_0 \text{ се отфрла} \mid H_0 \text{ е точна})$$
$$\beta = P(\text{грешка од тип II}) = P(H_0 \text{ се прифаќа} \mid H_0 \text{ не е точна})$$

- Веројатноста за грешка од тип I се нарекува и ***ниво на значајност на тестот***.

Моќ на тест

- Величината **1 – *β*** се нарекува ***моќ на тестот***, и одговара на веројатноста да се отфрли нултата хипотеза кога истата не е точна.

- Моќта на тестот всушност е мерка за **сензитивноста** на тестот, при што под сензитивност се подразбира способност на тестот да ги открие разликите.

Пример 1

- Познато е дека во една кутија се наоѓаат 10 топчиња, но не е сигурно дали се 6 црвени и 4 зелени или се 5 црвени и 5 зелени. Затоа се дефинираат следните хипотези:

  H_0: во кутијата има 6 црвени и 4 зелени;

  H_a: во кутијата има 5 црвени и 5 зелени топчиња.

- Се извлекува едно топче од кутијата и се дефинира следното правило за тестирање:
  - хипотезата H_0 не се отфрла ако и само ако се извлече црвено топче;
  - во спротивно H_0 се отфрла.
- Ќе ги определиме веројатностите на грешка од прв и втор тип.

Пример 1 - продолжение

- Грешка од прв тип ќе се појави, ако се отфрли нултата хипотеза кога таа е точна. Ако H_0 е точна, тогаш во кутијата има 6 црвени и 4 зелени топчиња. Хипотезата ќе ја отфрлиме, ако се извлече зелено топче, а веројатноста тоа да се случи е 4/10 = 0.4. Оттука,

$$\alpha = P\{\text{извлечено е зелено топче} \mid H_0 \text{ е точна}\} = 0.4.$$

- Грешка од втор тип ќе се појави, ако нултата хипотеза се прифати, а не е точна, т.е. точна е алтернативната. Во тој случај, во кутијата има 5 црвени и 5 зелени топчиња. Нултата хипотеза ќе се прифати, ако се извлече црвено топче. Веројатноста ова да се случи, при овој состав на кутијата, е 5/10 = 0.5, значи

$$\beta = P\{\text{извлечено е црвено топче} \mid H_a \text{ е точна}\} = 0.5.$$

Грешка од I тип и грешка од II тип

- Да воочиме дека грешка од тип I и грешка од тип II не може да се појават истовремено.
  - Грешка од тип I се појавува само ако H_0 е точна;
  - Грешка од тип II се појавува само ако H_0 е неточна.
- Тестовите кои ќе ги користиме сакаме да имаат што е можно помала грешка од тип I и што е можно помала тип грешка од тип II.
- Меѓутоа, ако веројатноста на грешка од тип I опаѓа, тогаш веројатноста на грешка од тип II расте, па не може истовремено да се минимизираат веројатностите за двата типа на грешка.

- Затоа, при тестирање на хипотезите однапред се задава веројатноста за грешка од тип I или нивото на значајност на тестот, а потоа се бараат тестови кои имаат најмала можна грешка од тип II, односно најголема моќ.

Избор на α

- Најчесто, се избира $\alpha = 0.05$ или $\alpha = 0.01$.
- Изборот на α, т.е. изборот на веројатноста да ја отфрлиме нултата хипотеза кога таа е точна, го покажува следното:
  - Нека тестот се повторува повеќе пати.
  - Секое повторување на тестот ќе го сметаме за обид да се најде доказ дека нултата хипотеза е неточна.
  - Ако $\alpha = 0.05$, тогаш може да се смета дека таков доказ ќе се најде во најмногу 5% од обидите, ако нултата хипотеза е точна.
  - Ако, пак, $\alpha = 0.01$, тоа значи дека при повторување на тестот повеќе пати, нултата хипотеза нема да се прифати во најмногу 1% од случаите, ако таа е точна.

Статистички тест. Тест статистика

Дефиниција 3. Секое правило, според кое при секоја реализација $\boldsymbol{x}$ на случајниот примерок $\boldsymbol{X}$ хипотезата H_0 или не ја отфрламе или ја отфрламе се нарекува ***статистички тест***.

- Правилата за тестирање или статистичките тестови ќе ги засноваме на вредностите на одредени статистики дефинирани на случајниот примерок кои се нарекуваат тест статистики.
- Следен чекор во постапката на тестирање на нултата хипотеза е изборот на ***тест статистика***. Тој избор зависи од хипотезата која ја тестираме.
  - Најчесто се поаѓа од оценувач на параметарот кој има одредени оптимални својства.
  - Ако се тестираат хипотези за математичкото очекување и дисперзијата на статистичкото обележје најчесто се поаѓа од некој точкаст оценувач на параметарот.

Да се потсетиме...

Непознат параметар	Точкаст оценувач
Математичко очекување (μ)	$\bar{X}$
Пропорција/веројатност (p)	$\hat{P}$
Дисперзија (σ^2)	S^2
Разлика на мат. очекувања на две обележја ($\mu_1 - \mu_2$)	$\bar{X}_1 - \bar{X}_2$

Критичен домен

- Претпоследен чекор во еден статистички тест е определувањето на ***критичниот домен***.
- Тоа е всушност, област на отфрлање на нултата хипотеза. Имено, реалната оска е поделена на две подобласти: област на прифаќање на нултата хипотеза и област на нејзино отфрлање (критичен домен).
    - Критичниот домен ќе го означуваме со C и тој се избира така што веројатноста тест статистиката (на пример, Z) да прими вредност која припаѓа во тој критичен домен, ако нултата хипотеза е точна, е најмногу α, т.е.

$$P\{Z \in C \mid H_0 \text{ е точна}\} \leq \alpha.$$

    - Да воочиме дека последната веројатност е, всушност, веројатност на грешка од тип I, тоа е веројатноста да ја отфрлиме нултата хипотеза, ако таа е точна.
    - Критичниот домен C се бира така што таа веројатност треба да биде помала или еднаква на α.

Најмоќен критичен домен

- Како што е претходно кажано, најчесто се настојува за дадена веројатност на грешка од прв тип α, да се определи минималната вредност на веројатноста за грешката од втор тип β.

- Критичниот домен за кој тоа е исполнето се нарекува ***оптимален критичен домен*** или ***најмоќен критичен домен***.

- Постојат методи за определување на најмоќен критичен домен за тестирање на една хипотеза, но овде нема да се задржуваме на објаснување на тие методи.

Донесување на одлука

- Во последниот чекор од процесот на тестирање, се пресметува вредноста на тест статистика врз основа на даден примерок и се проверува дали таа припаѓа на определениот критичен домен C.
  - Ако таа вредност припаѓа на C, тогаш нултата хипотеза се отфрла како неточна,
  - во спротивно нултата хипотеза не се отфрла, т.е. се прифаќа како точна.
- Ако податоците од случајниот примерок се такви што се отфрла H_0 велиме дека разликата меѓу податоците и нултата хипотеза е ***значајна*** или ***тестот е значаен***, или ***сигнификантен***.
- Доколку пак според случајниот примерок, нултата хипотеза H_0 не се отфрла велиме дека податоците ја поддржуваат нултата хипотеза или тестот не е значаен односно не е сигнификантен.

Алгоритам за тестирање на хипотези

1. Од контектот на задачата да се идентификува параметарот од интерес.
2. Се поставува нултата хипотеза, H_0 и соодветна алтернативна хипотеза H_a.
3. Се избира нивото на значајност на тестот α.
4. Се определува соодветна тест статистика.
5. Се определува критичниот домен C за избраната статистика.
6. Се пресметува вредноста на тест статистиката.
7. Се проверува дали пресметаната вредност на тест статистиката припаѓа на критичниот домен C или не. Ако припаѓа, тогаш нултата хипотеза се отфрла како неточна, во спротивно, таа се прифаќа како точна.

ТЕСТОВИ ЗА МАТЕМАТИЧКОТО ОЧЕКУВАЊЕ

Тестови за математичко очекување на статистичко обележје

- Нека $(X_1,\ldots,X_n)$ е случаен примерок за обележјето X чие математичко очекување е непознато.

- Ќе разгледаме тестови за математичко очекување на обележјето X, и тоа во два случаи:
  - Обележјето има нормална распределба и произволна големина на примерок.
  - Обележјето нема нормална распределба, но даден е голем случаен примерок од популацијата.

- Нека нивото на значајност на тестот е α.

Тестови за математичко очекување кај нормална распределба

- Понатаму ќе разгледаме еднострани и двострани тестови за математичкото очекување на обележје X кое има нормална распределба $N(\mu, \sigma^2)$ при неколку сценарија:

  1. познато σ^2, произволна големина на примерок

  2. непознато σ^2

     2.1 мал примерок

     2.2 голем примерок

ТЕСТОВИ ЗА МАТЕМАТИЧКОТО ОЧЕКУВАЊЕ НА ОБЕЛЕЖЈЕ $X \sim N(\mu, \sigma^2)$ ПРИ ПОЗНАТО σ^2 И ПРОИЗВОЛНА ГОЛЕМИНА НА ПРИМЕРОК

1. $X \sim N(\mu, \sigma^2)$, σ^2 познато

- Нека се тестираат хипотезите:

$$H_0: \mu = \mu_0$$
$$H_a: \mu \neq \mu_0$$

- Во овој случај, имаме двостран тест.
- Како да се избере тест статистиката во овој случај?
- Тргнуваме од точкастиот оценувачот за μ. Тоа е $\bar{X}_n$, т.е. просекот на примерокот. Кога σ^2 е познато, $\bar{X}$ има нормална распределба $N(\mu, \sigma^2/n)$. При претпоставка дека нултата хипотеза е точна, статистиката

$$Z_0 = \frac{\bar{X}_n - \mu_0}{\sigma / \sqrt{n}} = \frac{\bar{X}_n - \mu_0}{\sigma}\sqrt{n}$$

има нормална $N(0,1)$ распределба.

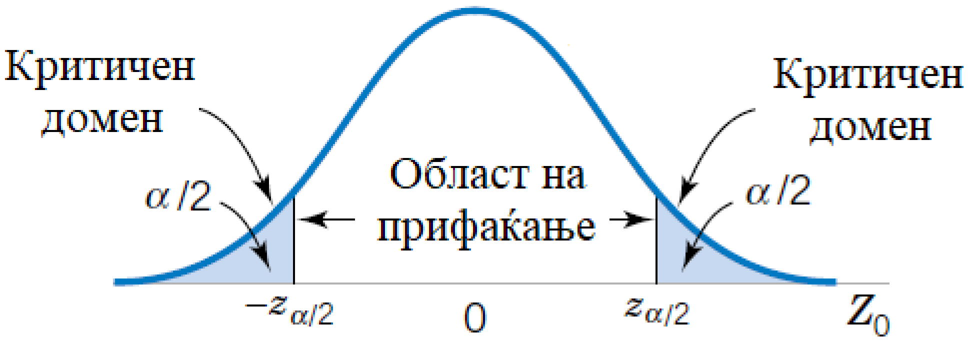
Критичен домен

- Критичниот домен се избира така што

$$P\{Z \in C \mid H_0 \text{ е точна}\} = \alpha,$$

и тој има облик:

$$C = (-\infty, -z_{\alpha/2}) \cup (z_{\alpha/2}, +\infty).$$

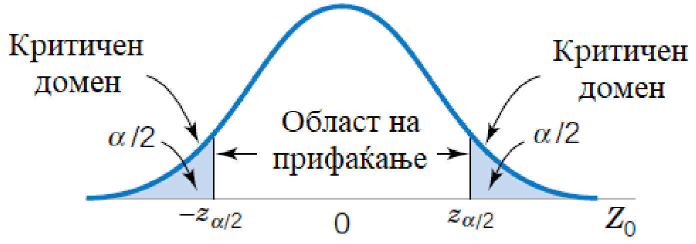


- Вредноста $z_{\alpha/2}$ се чита од таблица за нормална нормирана распределба, така што $\Phi(z_{\alpha/2})=1-\alpha/2$.

Критичен домен

- Како што претходно спомнавме, овде нема да се задржуваме на објаснување на методите за определување на најмоќен критичен домен за тестирање на една хипотеза.
- Но, сепак да појасниме зошто критичниот домен е избран на ваков начин? Зошто од краевите на распределбата?
- Да воочиме дека тест статистиката

$$Z_0 = \frac{\bar{X}_n - \mu_0}{\sigma}\sqrt{n}$$

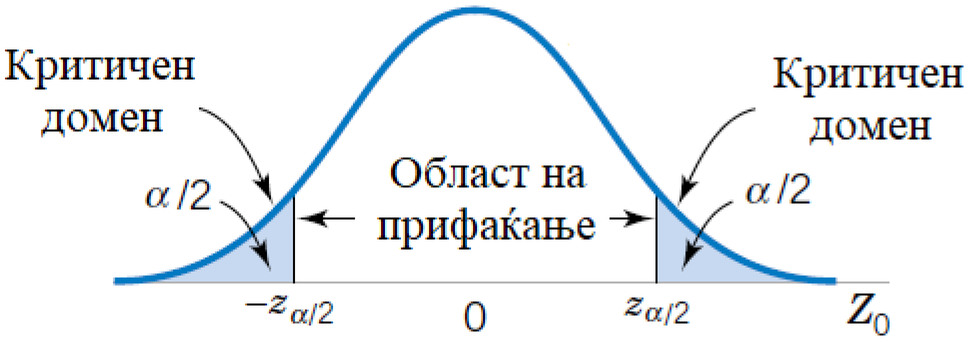


на некој начин, ја мери разликата помеѓу $\bar{X}_n$ (оценетата вредност на математичкото очекување според дадениот примерок) и μ_0 (претпоставената вредност на непознатото математичко очекување). Разликата меѓу нив ќе биде поголема, кога Z_0 прима вредности подалеку од 0.

- Алтернативата хипотеза е H_a: $\mu \neq \mu_0$, па не интересираат вредностите од двете страни на 0.

Критичен домен и интервал на доверба

- Да воочиме дека постои тесна врска помеѓу областа на прифаќање на нултата хипотеза (H_0: $\mu = \mu_0$) и интервал на доверба за непознатото математичко очекување μ, со ниво на доверба $1 - \alpha$.



- Имено, областа на прифаќање е определена со интервалот ($-z_{\alpha/2}$, $z_{\alpha/2}$) во кој треба да припаѓа вредноста на Z-статистиката која се користи за определување на интервалот на доверба.
- Притоа, ако вредноста на параметарот μ_0 одредена со нултата хипотеза е содржана во 95% интервал на доверба, тогаш нултата хипотеза не се отфрла на ниво на значајност 0,05.
- Ако вредноста одредена со нултата хипотеза не е во интервалот, нултата хипотеза ќе се отфрли на ниво 0,05.

Пример 2

Познато е дека очекуваниот крвен притисок кај мажите од 35 до 44 годишна возраст изнесува 127 единици со стандардна девијација од 7 единици. При систематски преглед на 70 случајно избрани вработени во една компанија, утврдено е дека нивниот просечен крвен притисок е 125. Дали при ниво на значајност $\alpha = 0.05$, може да се заклучи дека очекуваниот крвен притисок на вработените во таа компанија отстапува од стандардите? Што може да се заклучи, ако нивото на значајност е $\alpha = 0.01$? Се претпоставува дека висината на крвниот притисок има нормална распределба.

Решение: Нека X е обележјето – крвен притисок кај мажите од 35 до 44 годишна возраст и нека μ е математичкото очекување на оваа случајна променлива.

Се тестираат хипотезите:

$$H_0 \colon EX = 127$$

$$H_a \colon EX \neq 127$$

Пример 2 - продолжение

$$H_0: EX = 127$$
$$H_a: EX \neq 127$$

- Се пресметува вредноста на тест-статистиката Z_0. Се добива:

$$z_0 = \frac{\bar{x}_n - \mu_0}{\sigma}\sqrt{n} = \frac{125-127}{7}\sqrt{70} = -2.39.$$

- Од таблица за $N(0,1)$ распределба, се чита $z_{\alpha/2} = z_{0.025}$ така што $\Phi(z_{0.025}) = 1 - \alpha/2 = 0.975$. Наоѓаме $z_{0.025} = 1.96$.
- Значи, критичниот домен е $C = (-\infty, -1.96) \cup (1.96, +\infty)$.
- Вредноста на тест статистиката $z_0 = -2.39 \in C$, па H_0 се отфрла.
- Заклучуваме дека со ниво на значајност $\alpha = 0.05$, очекуваниот крвен притисок на вработените во компанијата отстапува од стандардите.

Пример 2 - продолжение

- Ако $\alpha = 0.01$, тогаш од таблица за $N(0,1)$ распределба, се чита $z_{\alpha/2} = z_{0.005}$ така што $\Phi(z_{0.005}) = 1 - \alpha/2 = 0.995$, па $z_{0.005} = 2.58$. За критичниот домен во овој случај, се добива

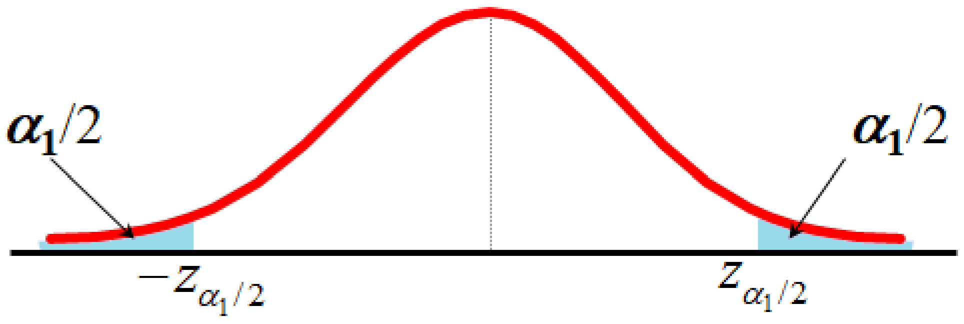
$$C = (-\infty, -2.58) \cup (2.58, +\infty).$$

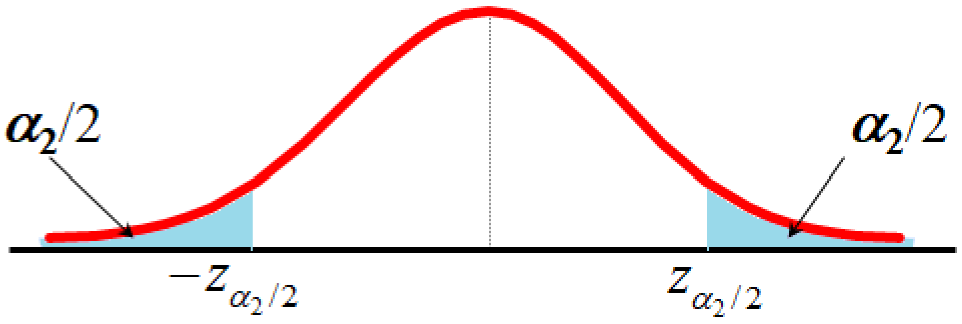
- Вредноста на тест статистиката $z_0 = -2.39 \notin C$, па во овој случај, H_0 не се отфрла.

- Ако означиме $C_1 = (-\infty, -1.96) \cup (1.96, +\infty)$ и $C_2 = (-\infty, -2.58) \cup (2.58, +\infty)$, јасно е дека $C_1 \supseteq C_2$.
- Имено, со намалување на нивото на значајност се намалува и критичниот домен.

Врска помеѓу критични домени со различни нивоа на значајност

- Во општ случај, нека за две нивоа на значајност α_1 и α_2 важи $\alpha_1 < \alpha_2$ и нека C_1 и C_2 се соодветните критични домени. Тогаш важи дека $C_1 \subseteq C_2$.
- Затоа, ако нултата хипотеза се отфрли со ниво на значајност α_1, т.е. ако вредноста на тест статистиката припаѓа на C_1, таа вредност ќе припаѓа и на C_2, па нултата хипотеза се отфрли и со ниво на значајност α_2.
- Имено, нултата хипотеза ќе се отфрли со секое ниво на значајност што е поголемо од α_1.



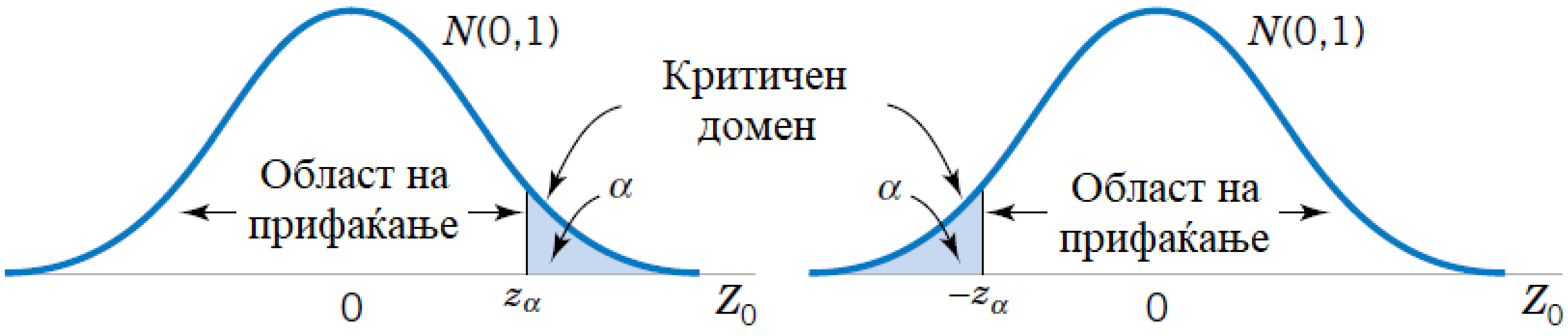


Еднострани тестови

- Во продолжение, ќе ги разгледаме едностраните тестови

$$H_0: \mu = \mu_0 \qquad\qquad H_0: \mu = \mu_0$$
$$H_a: \mu > \mu_0 \qquad \text{или} \qquad H_a: \mu < \mu_0$$

- За тестирање на овие хипотези се користи истата статистика Z_0, но сега се менува критичниот домен.
- Во првиот случај, критичниот домен е $C = (z_\alpha, +\infty)$,
- Во вториот случај, тој е $C = (-\infty, -z_\alpha)$.
- Притоа, z_α е онаа вредност од $N(0,1)$ распределбата, за која $\Phi(z_\alpha) = 1 - \alpha$.



ТЕСТОВИ ЗА МАТЕМАТИЧКОТО ОЧЕКУВАЊЕ НА ОБЕЛЕЖЈЕ $X \sim N(\mu, \sigma^2)$ ПРИ НЕПОЗНАТО σ^2 И МАЛ ПРИМЕРОК

$X \sim N(\mu, \sigma^2)$, σ^2 непознато – мал примерок

- Нека се тестираат хипотезите:

$$H_0: \mu = \mu_0 \qquad\qquad H_0: \mu = \mu_0 \qquad\qquad H_0: \mu = \mu_0$$
$$H_a: \mu \neq \mu_0 \quad \text{или} \quad H_a: \mu > \mu_0 \quad \text{или} \quad H_a: \mu < \mu_0$$

(двостран тест)

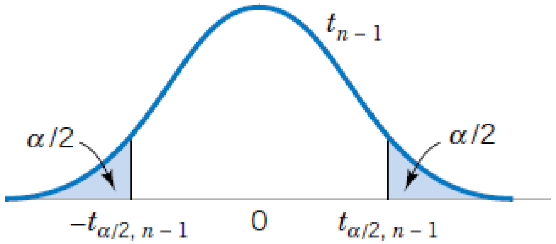
- Слично како и претходно, тргнуваме од точкастиот оценувач за μ, а тоа е $\bar{X}_n$, т.е. просекот на примерокот.
- Кога дисперзијата на обележјето σ^2 е непозната и имаме мал примерок, тогаш ако е точна нултата хипотеза, статистиката

$$T_0 = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{\bar{X} - \mu_0}{S}\sqrt{n}$$

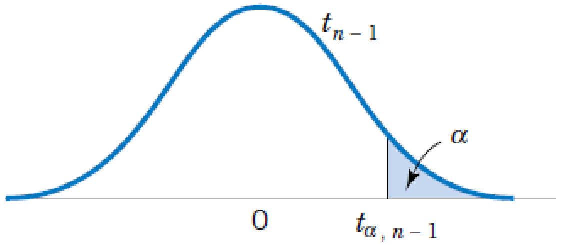
има t-распределба со $n - 1$ степени на слобода.

$X \sim N(\mu, \sigma^2)$, σ^2 непознато – мал примерок

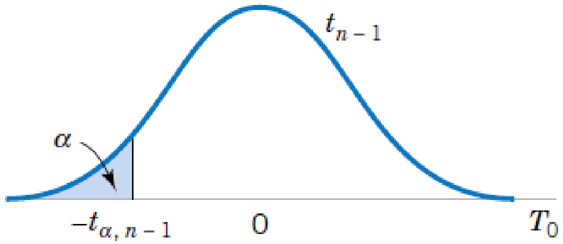
- Исто како и во претходниот случај, критичниот домен зависи од алтернативната хипотеза.
  - Ако имаме двостран тест, т.е. H_a: $\mu \neq \mu_0$, тогаш критичниот домен е
$$C = (-\infty, -t_{\alpha/2,\, n-1}) \cup (t_{\alpha/2,\, n-1}, +\infty),$$
каде $t_{\alpha/2,\, n-1}$ се чита од таблица на t – распределба, така што $P\{T > t_{\alpha/2,\, n-1}\} = \alpha/2$.
  - Ако алтернативната хипотеза е H_a: $\mu > \mu_0$, тогаш критичниот домен е $C = (t_{\alpha,\, n-1}, +\infty)$, каде $t_{\alpha,\, n-1}$ се чита од таблица на t – распределба, така што $P\{T > t_{\alpha,\, n-1}\} = \alpha$.
  - Ако пак, алтернативната хипотеза е H_a: $\mu < \mu_0$, тогаш критичниот домен е $C = (-\infty, -t_{\alpha,\, n-1})$, каде $t_{\alpha,\, n-1}$ се чита од таблица на t – распределба, така што $P\{T > t_{\alpha,\, n-1}\} = \alpha$.



H_a: $\mu \neq \mu_0$



H_a: $\mu > \mu_0$



H_a: $\mu < \mu_0$

Пример 3

Познато е дека потрошувачката на гориво на автомобили од типот Пежо има нормална распределба со очекувана вредност од 8 литри на 100 километри. Од произведената серија случајно се избрани 17 автомобили и добиена е просечна потрошувачка од 8.8 литри на 100 километри со стандардна девијација 0.9. Дали со ниво на значајност $\alpha = 0.05$, може да се тврди дека серијата има поголема од предвидената потрошувачка?

Решение: Нека X е обележјето – потрошувачка на 100 km и нека μ е математичкото очекување на оваа случајна променлива.

Се тестираат хипотезите:

$$H_0: \mu = 8$$

$$H_a: \mu > 8$$

Пример 3 - продолжение

$$H_0 \colon \mu = 8$$

$$H_a \colon \mu > 8$$

- Се пресметува вредноста на тест-статистиката $t_{0.}$ Се добива:

$$t_0 = \frac{\overline{x}_n - \mu_0}{s}\sqrt{n} = \frac{8.8 - 8}{0.9}\sqrt{17} = 3.66.$$

- Од таблица за t-распределба, се чита $t_{\alpha,\, n-1} = t_{0,05;16} = 1.746$.
- Значи, критичниот домен е $C = (1.746, +\infty)$.
- Вредноста на тест статистиката $t_0 = 3.66 \in C$, па H_0 се отфрла.
- Заклучуваме дека потрошувачката е поголема од 8 литри на 100 km.

Пример 4

Се смета дека цената по ноќ во хотел во Скопје има нормална распределба со очекување 1680 ден. по ноќ. Земен е случаен примерок од 25 хотели и добиено е просек $\bar{x} = 1725$ и стандардна девијација $s = 154$. Да се тестира тврдењето за очекуваната цена со ниво на значајност $\alpha = 0.05$.

Решение: Нека X е обележјето – цената по ноќ во хотел во Скопје и нека μ е математичкото очекување на оваа случајна променлива.

Се тестираат хипотезите:

$$H_0: \mu = 1680$$

$$H_a: \mu \neq 1680$$

Пример 4 - продолжение

$$H_0 \colon \mu = 1680$$
$$H_a \colon \mu \neq 1680$$

- Се пресметува вредноста на тест-статистиката t_0. Се добива:

$$t_0 = \frac{\overline{x}_n - \mu_0}{s}\sqrt{n} = \frac{1725 - 1680}{154}\sqrt{25} = 1.46.$$

- Од таблица за t-распределба, се чита $t_{\alpha/2,\, n-1} = t_{0,025,\, 24} = 2.064$.
- Значи, критичниот домен е $C = (-\infty, 2{,}064) \cup (2.064, +\infty)$.
- Вредноста на тест статистиката $t_0 = 1.46 \notin C$, па H_0 не се отфрла.
- Заклучуваме дека очекуваната цена по ноќ во хотел во Скопје изнесува 1680 денари.

ТЕСТОВИ ЗА МАТЕМАТИЧКОТО ОЧЕКУВАЊЕ НА ОБЕЛЕЖЈЕ $X \sim N(\mu, \sigma^2)$ ПРИ НЕПОЗНАТО σ^2 И ГОЛЕМ ПРИМЕРОК

$X \sim N(\mu, \sigma^2)$, σ^2 непознато – голем примерок

- Кога е даден голем примерок ($n \geq 30$), тогаш t-распределбата тежи кон нормална нормирана распределба. Затоа, може да земеме дека

$$T_0 = \frac{\bar{X} - \mu_0}{S}\sqrt{n} \approx Z_0$$

и да работиме со соодветните z - тестови.

ТЕСТОВИ ЗА МАТЕМАТИЧКОТО ОЧЕКУВАЊЕ НА ОБЕЛЕЖЈЕ СО ПРОИЗВОЛНА РАСПРЕДЕЛБА И ГОЛЕМ ПРИМЕРОК

Тестови за математичко очекување на произволно обележје – голем примерок

- Честопати во пракса не можеме да претпоставиме нормална распределба на статистичкото обележје. Меѓутоа кога станува збор за голем примерок ($n \geq 30$), тогаш согласно централната гранична теорема за тестирање на хипотезите за математичкото очекување може да се користи Z – статистика.
- Ако дисперзијата на обележјето е позната, тогаш се користи статистиката

$$Z_0 = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n},$$

- Ако дисперзијата не е позната, тогаш се користи статистиката

$$Z_0 = \frac{\bar{X} - \mu_0}{S}\sqrt{n}.$$

- И во двата случаи, кога е точно H_0, Z_0 има асимптотски $N(0,1)$ распределба.

Пример 5

Една фабрика за лекови произведува антибиотици и очекуваната моќност на еден од тие антибиотици е 80%. Тестиран е случаен примерок од 100 капсули и добиена е просечна моќност од 79.7% и стандардна девијација од 2%. Дали произведените антибиотици се според пропишаните стандарди? Нека $\alpha = 0.01$.

Решение: Нека X е обележјето – моќност на антибиотикот и нека μ е математичкото очекување на оваа случајна променлива.

Се тестираат хипотезите:

$$H_0 : \mu = 80$$

$$H_a : \mu \neq 80$$

Пример 5 - продолжение

$$H_0 \colon \mu = 80$$

$$H_a \colon \mu \neq 80$$

- Се пресметува вредноста на тест-статистиката z_0. Се добива:

$$z_0 = \frac{\bar{x}_n - \mu_0}{s}\sqrt{n} = \frac{79.7 - 80}{2}\sqrt{100} = -1.5.$$

- Од таблица за z-распределба, се чита $z_{\alpha/2} = z_{0.005}$ така што $\Phi(z_{0.005}) = 1 - \alpha/2 = 0.995$. Наоѓаме $z_{0.005} = 2.58$.
- Значи, критичниот домен е $C = (-\infty, -2.58) \cup (2.58, +\infty)$.
- Вредноста на тест статистиката $z_0 = -1.5 \notin C$, па H_0 не се отфрла.
- Заклучуваме дека произведените антибиотици се според пропишаните стандарди.

Тест статистики за математичко очекување

- тест статистики за мат. очекување
  - сите распределби голем примерок $(n \geq 30)$
    - позната дисперзија: $Z_0 = \dfrac{\bar{X} - \mu_0}{\sigma}\sqrt{n}$
    - непозната дисперзија: $Z_0 = \dfrac{\bar{X} - \mu_0}{S}\sqrt{n}$
  - нормална распределба мал примерок
    - позната дисперзија: $Z_0 = \dfrac{\bar{X} - \mu_0}{\sigma}\sqrt{n}$
    - непозната дисперзија: $T_0 = \dfrac{\bar{X} - \mu_0}{S}\sqrt{n}$

p – вредност на тест

- Еден начин да се искажат резултатите од тестирањето на хипотезите е да се каже дали нултата хипотеза се отфрла или не за однапред дефинирано α - ниво на значајност на тестот.

- На пример, можеме да кажеме дека H_0: $\mu = 50$ е отфрлена со ниво на значајност 0.05.

- Овој заклучок често е несоодветен, бидејќи на тој што донесува одлука не му дава некаква идеја дали пресметаната вредност на тест-статистиката е на граница на критичната област или е многу далеку од оваа област.

p – вредност на тест

- За да се избегне ова честопати резултатите од тестирањето се соопштуваат преку таканаречена *p*-вредност на тест статистиката.

Дефиниција 4. *p*-вредност претставува најмалото ниво на значајност кое би водело до отфрлање на нултата хипотеза H_0 со дадените податоци.

- Ако вредноста на α е помала од добиената *p*-вредност, тогаш нултата хипотеза се прифаќа како точна. Во спротивно, се отфрла како неточна.
- *p*-вредноста носи повеќе информација за тежината на доказите против H_0 и затоа оној кој донесува одлуки може да изведе заклучок за било кое дадено ниво на значајност на тестот.

Пример 6

- Да ја разгледаме повторно задачата од Пример 5.
- Една фабрика за лекови произведува антибиотици и очекуваната моќност на еден од тие антибиотици е 80%. Тестиран е случаен примерок до 100 капсули и добиена е просечна моќност од 79.7% и стандардна девијација од 2%. Дали произведените антибиотици се според пропишаните стандарди? Да се определи p- вредност за овој тест. Што може да се заклучи ако $\alpha = 0.05$?

Решение: Нека X е обележјето – моќност на антибиотикот и нека μ е математичкото очекување на оваа случајна променлива.

Се тестираат хипотезите:

$$H_0\!: \mu = 80$$

$$H_a\!: \mu \neq 80$$

Пример 6 - продолжение

- Пресметаната вредност на тест-статистиката z_0 е:

$$z_0 = \frac{\overline{x}_n - \mu_0}{s}\sqrt{n} = \frac{79.7-80}{2}\sqrt{100} = -1.5.$$

- Се бара најмалото ниво на значајност кое би водело до отфрлање на нултата хипотеза H_0. Хипотезата ќе се отфрли, ако

$$|Z| > \left|\frac{\overline{x}_n - \mu_0}{s}\sqrt{n}\right| = \left|\frac{79.7-80}{2}\sqrt{100}\right| = 1.5,$$

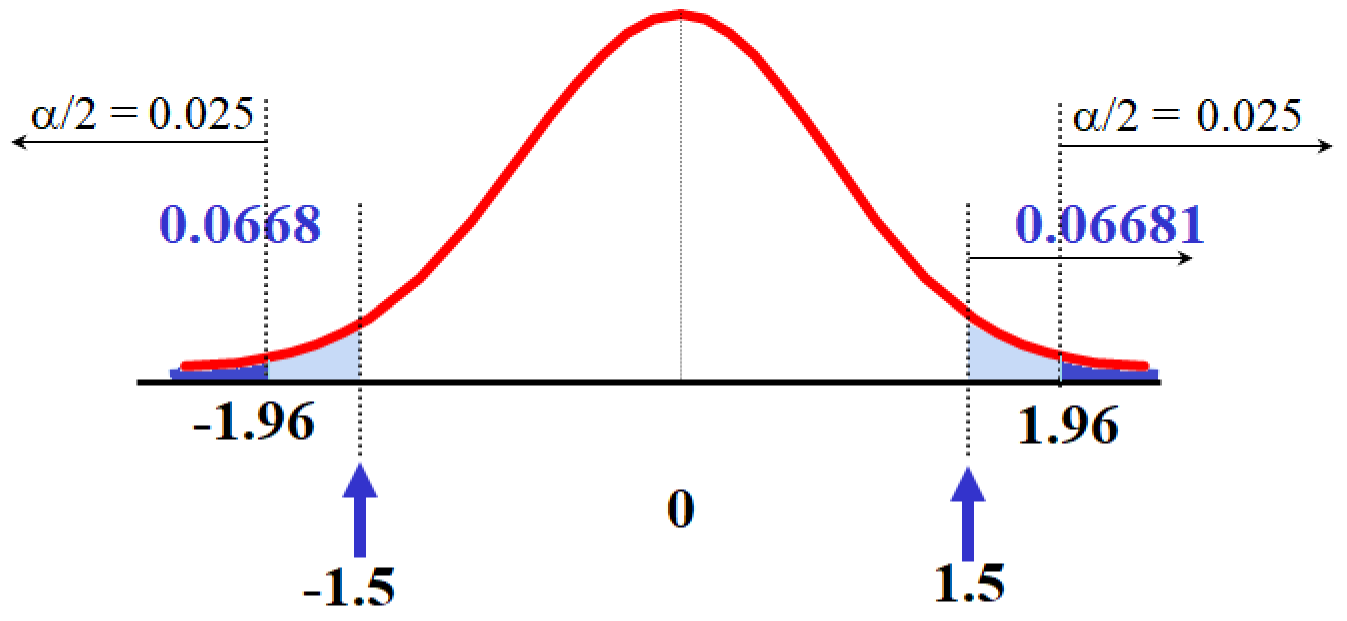
кога е точна H_0. Сега,

$$\begin{aligned} p - \text{вредност} &= P\{|Z| > 1.5 \mid H_0 \text{ е точна}\} \\ &= P\{Z \le -1.5 \mid H_0 \text{ е точна}\} + P\{Z \ge 1.5 \mid H_0 \text{ е точна}\} \\ &= \Phi(-1.5) + (1 - \Phi(1.5)) \\ &= 0.06681 + (1 - 0.93319) \\ &= 0.06681 + 0.06681 \\ &= 0.13362 \end{aligned}$$

Пример 6 - продолжение

$$p - \text{вредност} = P\{Z \leq -1.5 \mid H_0 \text{ е точна}\} + P\{Z \geq -1.5 \mid H_0 \text{ е точна}\}$$
$$= 0.06681 + 0.06681 = 0.13362$$

- p-вредност > α = 0.05, па следува дека хипотезата не се отфрла со ова ниво на значајност на тестот.



ТЕСТОВИ ЗА ВЕРОЈАТНОСТ НА НАСТАН,
(ТЕСТОВИ ЗА ПРОПОРЦИИ)
ГОЛЕМ ПРИМЕРОК

Тестови за веројатност на настан, (тестови за пропорции) голем примерок

- Нека p е веројатноста на настан A и таа е непозната. Треба да се спроведе постапка за тестирање дали таа веројатност прима некоја претпоставена вредност p_0.
- Нултата хипотеза ќе гласи:

$$H_0: p = p_0$$

наспроти една од алтернативните

$$H_a: p \neq p_0 \quad \text{или} \quad H_a: p > p_0 \quad \text{или} \quad H_a: p < p_0.$$

- За спроведување на тестирањето се зема случаен примерок $(X_1,\ldots,X_n)$ за обележјето X – индикатор на настанот A. Јасно е дека X_i прима вредност 0 (ако не се појави настанот A) или 1 (ако се појави настанот A).
- Видовме дека точкаст оценувач за параметарот p е $\hat{P} = \bar{X}$.
- Затоа, $\hat{P}$ може да се користи за добивање на тест статистика во овој случај.

Тестови за веројатност на настан, (тестови за пропорции) голем примерок

- Согласно централната гранична теорема кога станува збор за голем примерок, ако е точна нултата хипотеза статистиката

$$Z_0 = \frac{\hat{P} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$$

има асимптотски нормална $N(0,1)$ распределба.

- Затоа, критичниот домен се определува со користење на таблица за $N(0,1)$ распределба, и тој зависи од алтернативната хипотеза. Така,

$H_a: p \neq p_0:$ $\quad C = (-\infty, -z_{\alpha/2}) \cup (z_{\alpha/2}, +\infty)$

$H_a: p > p_0:$ $\quad C = (z_\alpha, +\infty)$

$H_a: p < p_0:$ $\quad C = (-\infty, -z_\alpha)$

Пример 7

Врз основа на претходни резултати познато е дека веројатноста еден стрелец да ја погоди целта е 0.9. После подготовките за натпревар, извршено е пробно гаѓање, во кое стрелецот ја погодил целта 92 пати од 100 гаѓања. Со ниво на значајност $\alpha = 0.01$, да се тестира хипотезата дека веројатноста за постигнување погодок не е променета, наспроти алтернативната дека е зголемена.

Решение: Означуваме

A: стрелецот ја погодил целта

и нека $p = P(A)$.

- Ги поставуваме следните хипотези:

$$H_0: p = 0.9$$

$$H_a: p > 0.9$$

Пример 7 - продолжение

$$H_0: p = 0.9$$
$$H_a: p > 0.9$$

- Врз основа на добиениот примерок се пресметува $\hat{p} = \bar{x} = 0.92$.
- За вредноста на тест-статистиката z_0 се добива:

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.92 - 0.9}{\sqrt{0.09/100}} = 0.67$$

- Критичниот домен е $C = (z_{0.01}, +\infty) = (2.33, +\infty)$.
- Вредноста на тест статистиката $z_0 = 0.67 \notin C$, па H_0 не се отфрла.
- Заклучуваме дека веројатноста за постигнување на погодок не е зголемена.

Пример 8

Една маркетинг компанија тврди дека добива одговор на 8% од испратените анкети. За да се тестира ова испратени се 500 анкетни прашалници и добиени се 25 одговори.

Да се тестира тврдењето со ниво на значајност $\alpha = 0.05$.

Решение: Означуваме

A: на анкетниот прашалник е добиен одговор

и нека $p = P(A)$.

- Ги поставуваме следните хипотези:

$$H_0: p = 0.08$$
$$H_a: p \neq 0.08$$

Пример 8 - продолжение

$$H_0: p = 0.08$$
$$H_a: p \neq 0.08$$

- Врз основа на добиениот примерок се пресметува $\hat{p} = \bar{x} = 0.05$.
- За вредноста на тест-статистиката z_0 се добива:

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.05 - 0.08}{\sqrt{0.08 \cdot 0.92 / 500}} = -2.47.$$

- Критичниот домен е $C = (-\infty, z_{0.025}) \cup (z_{0.025}, +\infty) = (-\infty, -1.96) \cup (1.96, +\infty)$.
- Вредноста на тест статистиката $z_0 = -2.47 \in C$, па H_0 се отфрла.
- Заклучуваме дека веројатноста за одговор на анкетата е различна од 8%.

ТЕСТОВИ ЗА ДИСПЕРЗИЈА (И СТАНДАРДНА ДЕВИЈАЦИЈА) НА ОБЕЛЕЖЈЕ СО НОРМАЛНА РАСПРЕДЕЛБА

Тестови за дисперзијата и стандардната девијација на нормална распределба

- Нека обележјето X има $N(\mu,\sigma^2)$ и притоа и μ и σ се непознати.
- Сакаме да ја тестираме хипотезата

$$H_0 : \sigma = \sigma_0$$

наспроти една од алтернативните хипотези

$$H_a\!: \sigma \neq \sigma_0 \qquad \text{или} \qquad H_a\!: \sigma > \sigma_0 \qquad \text{или} \qquad H_a\!: \sigma < \sigma_0$$

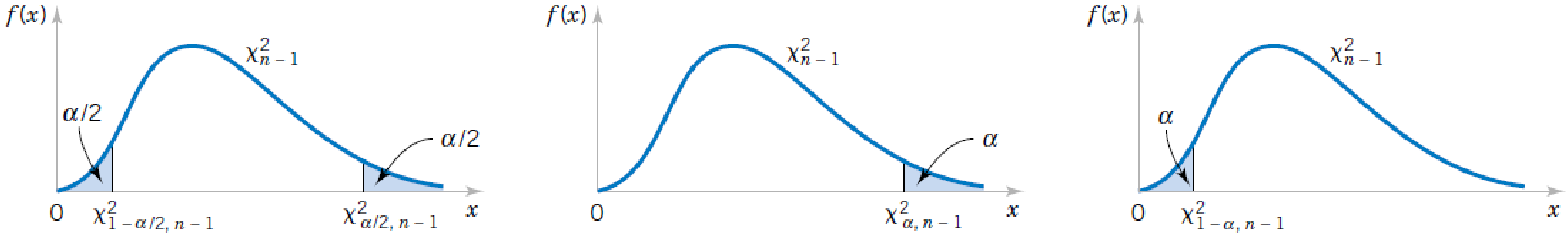
- За тестирање ќе ја користиме тест статистиката

$$\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2},$$

која кога е точна нултата хипотеза има χ^2 - распределба со $n-1$ степени на слобода.

Тестови за дисперзијата и стандардната девијација на нормална распределба

- Притоа, исто како и кај претходните тестови, критичниот домен зависи од алтернативната хипотеза.
  - За H_a: $\sigma \neq \sigma_0$, критичниот домен е
  $$C = (0, \chi^2_{1-\alpha/2,n-1}) \cup (\chi^2_{\alpha/2,n-1}, +\infty).$$
  - За H_a: $\sigma > \sigma_0$, критичниот домен е $C = (\chi^2_{\alpha,n-1}, +\infty)$.
  - За H_a: $\sigma < \sigma_0$, критичниот домен е $C = (0, \chi^2_{1-\alpha,n-1})$.



H_a: $\sigma \neq \sigma_0$ H_a: $\sigma > \sigma_0$ H_a: $\sigma < \sigma_0$

Пример 9

Врз основа на случаен примерок со обем $n = 25$ за нормално распределено обележје X, добиена е дисперзија 22. Дали со ниво на значајност $\alpha = 0.05$, може да се заклучи дека дисперзијата на обележјето е поголема од 15?

Решение: Ги поставуваме хипотезите:

$$H_0 : DX = 15$$

$$H_a : DX > 15$$

- Вредноста на тест статистиката е

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{24 \cdot 22}{15} = 35.2.$$

- За $\alpha = 0.05$, критичниот домен е $C = (\chi^2_{0.05,24}, +\infty) = (36.41, +\infty)$.
- $\chi_0^2 \notin C$, па нултата хипотеза не се отфрла.

Критичниот домен зависи само од алтернативната хипотеза

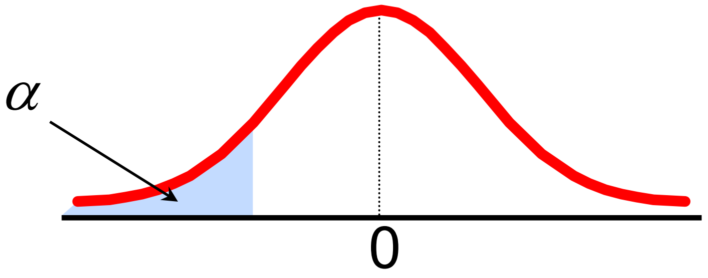
- Сите тестови што ги разгледавме досега беа тестови во кои нултата хипотеза е проста, а алтернативната е сложена.
- Она што е битно да се напомене дека ако алтернативната хипотеза остане иста, а нултата стане сложена, тогаш критичниот домен нема да се промени.
- Она што е битно да се внимава е дека множествата вредности на параметарот во нултата и во алтернативната хипотеза мора да се дисјунктни, за да отфрлањето на едната хипотеза би значело прифаќање на другата.

Критичен домен

- Ако се спроведува тест за вредност на математичко очекување, тогаш ги имаме следните случаи.

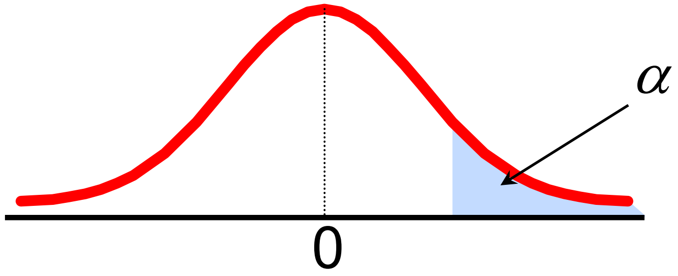
$H_0: \mu \geq \mu_0$ (или $\mu = \mu_0$)

$H_A: \mu < \mu_0$



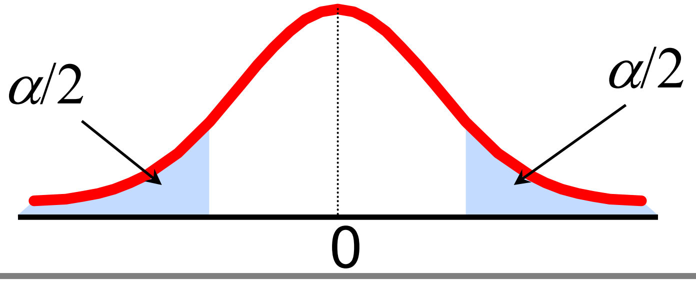
$H_0: \mu \leq \mu_0$ (или $\mu = \mu_0$)

$H_A: \mu > \mu_0$



$H_0: \mu = \mu_0$

$H_A: \mu \neq \mu_0$



- Исто поставување на хипотезите може да се спроведе и во сите останати случаи.

Пример 10

Направено е истражување за просечната плата на ИТ инженери. Земен е примерок од 50 ИТ инженери и добиено е дека нивната просечна плата е 62 570 денари со стандардна девијација од 7 500 денари. Со ниво на значајност 0.05, да се тестира хипотеза дека просечната плата е најмалку 61 500 денари наспроти алтернативната дека е помала 61 500 денари.

Решение: Ги поставуваме хипотезите:

$$H_0 \colon EX \geq 61\,500$$

$$H_a \colon EX < 61\,500$$

- Имаме голем примерок, па користиме Z – статистика.

$$z_0 = \frac{\bar{x} - \mu_0}{s}\sqrt{n} = \frac{62570 - 61500}{7500}\sqrt{50} = 1.17.$$

- Критичниот домен е $C = (-\infty, -z_{0.05}) = (-\infty, -1.65)$.
- $z_0 \notin C$, па нултата хипотеза не се отфрла, т.е. просечната плата на ИТ инженерите е најмалку 61 500 денари.

Пример 11

- Заради проблемите со јавниот транспорт и густиот сообраќај, работниците во еден голем град, тешко може да стигнат точно на време на работа и да ги исполнат своите работни обврски до крајот на работното време. За таа цел една компанија одлучува да воведе флексибилно работно време, според кое секој вработен ги распоредува часовите кои треба да ги одработи. Според проценките на компанијата за нејзино ефикасно работење, потребно е секој вработен да работи просечно најмалку по 8 часа дневно. Случајно се избрани 70 вработени и тие требало да испратат пробен распоред на нивното флексибилно работно време. Утврдено е дека просечниот број работни часови во понеделник е 7.7 со стандардна девијација од 2.7 часа. Дали врз основа на овој примерок, со ниво на значајност на тестот $\alpha = 0.05$, може да се заклучи дека бројот на работни часови во понеделник на сите вработени во оваа компанија ќе биде помал од 8? Задачата да се реши преку p-вредност.

Пример 11 - продолжение

- Нека X е обележјето – број на работни часови во понеделник на еден вработен и нека μ е математичкото очекување на оваа случајна променлива.
- Се тестираат хипотезите:

$$H_0\!: \mu \geq 8$$

$$H_a\!: \mu < 8$$

- Од условите на задачата, дадено е дека $n = 70$, $\bar{x} = 7.7$, $s = 2.7$.
- Вредноста на Z статистиката (имаме голем примерок) е

$$z_0 = \frac{\bar{x} - \mu_0}{s}\sqrt{n} = \frac{7.7 - 8}{2.7}\sqrt{70} = -0.93.$$

Пример 11 - продолжение

$$z_0 = \frac{\bar{x} - \mu_0}{s}\sqrt{n} = \frac{7.7 - 8}{2.7}\sqrt{70} = -0.93.$$

- Алтернативната хипотеза е H_a: $\mu < 8$, па

$$p\text{-вредност} = P\{Z < -0.93\} = 0.17619.$$

- Бидејќи, p-вредност $= 0.17619 > 0.05 = \alpha$, заклучуваме дека нултата хипотеза не се отфрла.

- Тоа значи дека очекуваниот број работни часови на еден вработен во произволно избран понеделник ќе биде најмалку 8.