Упражнение 8

Проверка на хипотези.
Тестове за проверка на хипотези за една средна стойност (критична област).

Алгоритъм за тестване на хипотези

критична област

1. Напишете нулевата H_0 и алтернативната H_1 хипотеза

2. Определете нивото на значимост α

3. Определете статистиката и извадковото разпр.

4. Получете критичната област

5. Направете извод

6. Интерпретация на извода

Проверка на хипотези за средна стойност

ПЪРВИ НАЧИН: КРИТИЧНА ОБЛАСТ

Контрахипотеза за популационното средно	Статистика при известна дисперсия σ^2	Критична област
$H_1: \mu>\mu_0$	$z=\dfrac{\overline{x}-\mu_0}{\dfrac{\sigma}{\sqrt{n}}}\in N(0,1)$	$\left(z_{1-\alpha},\infty\right)$
$H_1: \mu<\mu_0$		$\left(-\infty,-z_{1-\alpha}\right)$
$H_1: \mu\neq\mu_0$		$\left(-\infty,-z_{1-\frac{\alpha}{2}}\right)\cup\left(z_{1-\frac{\alpha}{2}},\infty\right)$

Извадката е направена от нормална популация с
- **известна дисперсия**

 или
- **неизвестна дисперсия, но извадка с голям обем, n>30 (В този случай във формулата популационното стандартно отклонение σ се замества с неговата точкова оценка s)**

Проверка на хипотези за средна стойност

Контрахипотеза за популационното средно	Статистика при НЕизвестна дисперсия σ^2	Критична област
H_1: $\mu>\mu_0$	$t=\dfrac{(\bar{x}-\mu_0)}{\dfrac{s}{\sqrt{n}}}\in t(n-1)$	$\left(t_{1-\alpha}(n-1),\infty\right)$
H_1: $\mu<\mu_0$		$\left(-\infty,-t_{1-\alpha}(n-1)\right)$
H_1: $\mu\neq\mu_0$		$\left(-\infty,t_{1-\frac{\alpha}{2}}(n-1)\right)\cup\left(t_{1-\frac{\alpha}{2}}(n-1),\infty\right)$

Извадката с малък обем n<30 е направена от нормална популация с неизвестна дисперсия
Или неизвестно популационно стандартно отклонение

13.4.За да се определи съдържанието на бактерии във водата на голямо езеро, за което се знае че е нормално разпределено, се вземат 50 проби от по 100 милилитра вода от различни места на брега и в лабораторията се измерва количеството бактерии в пробите. Намерено е, че средното количество бактерии е 11,95 (в стотици) и стандартно отклонение 8 (в стотици). Има ли статистически значимо основание да се смята, че броят бактерии в 100 милилитра от водата в това езеро е повече от 1000?

Решение: Интерпретация а данните:
50=n обем на извадката
8= извадково стандартно отклонение s(статистика)
10= константата от хипотезата
11.95=$\bar{x}$ извадковото средно (статистика)
Избираме ниво на значимост 0.05=α

дали броят бактерии в това езеро е повече от 10(стотици) => едностранен тест

H_0: μ=10 H_1: μ>10 Използваме z-статистиката(защо)

Z=$\frac{11.95-10}{8}\sqrt{50} = 1.72357$

При ниво на значимост 0.05 се търси точка от Приложение 1 надясо от която лицето е 0.05 или наляво от която лицето е 1-α=0.95 , това е точката 1.64 и КО е $(1.64, \infty) \Rightarrow$
Извод: 1.72357 е вътре в КО ⇒ Отхвърляме H_0. Броят бактерии в 100 милилитра от водата в това езеро е повече от 1000.

13.6. Производител пакетира течен сапун в бутилки с тегло 500 мл. За да се провери дали машината за пълнене е регулирана добре се избират случайно 100 бутилки и се претеглят. Намерено е, че средното тегло е 498.5 мл със стандартно отклонение 5 мл. Дават ли измерванията достатъчно основание да се настоява за пренастройка на машината?

Решение: Интерпретация а данните:

5= извадково стандартно отклонение s (статистика)

500= константата от хипотезата

498.5=$\overline{\mathrm{x}}$ извадковото средно (статистика)

100=n=обем на извадката

Избираме ниво да значимост 0.05=α

дали е регулирана добре => двустранен тест

H_0: μ=500 H_1: μ≠500

Неизвестна популационна дисперсия, но голям обем. Използваме $z = \dfrac{\overline{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \in N(0,1)$

Z=$\dfrac{498.5-500}{5}\sqrt{100} = -3$

При ниво на значимост 0.05 търсим точка, надясно от която лицето е α/2=0.025, или наляво от която лицето е 1-α/2=0.975), това е 1.96. Тогава КО е

$(-\infty, -1.96)$и$(1.96, \infty) \Rightarrow$ Извод −3 е вътре в КО ⇒ Отхвърляме H_0. Машината за пълнене НЕ е регулирана добре.

==

13.7. Производител на лютеница я пакетира в буркани с етикети, на които е записано *нето 400 г.* Известно е, че теглото е нормално разпределено със стандартно отклонение 10 грама. Направена е извадка от 16 буркана и е получено, че тяхното средно тегло е 404 г.
При ниво на значимост 0,05, трябва ли да се препоръча регулиране на машината?

Решение: Интерпретация а данните:
10= популационно стандартно отклонение σ(параметър)
400= константата от хипотезата
404=$\overline{x}$ извадковото средно (статистика)
16=n=обем на извадката
ниво да значимост 0.05=α

дали е регулирана добре => двустранен тест

H_0: μ=400 H_1: μ≠400

Известна популационна дисперсия. Използваме статистиката $z = \dfrac{\overline{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \in N(0,1)$

Z=$\frac{404-400}{10}\sqrt{16} = 1.6$

При ниво на значимост 0.05 се търси 1-α/2=0.975 и КО е $(-\infty, -1.96)$ и $(1.96, \infty) \Rightarrow$ Извод 1.6 НЕ е вътре в КО ⇒ Няма основание да отхвърлим H_0. Няма основание да се препоръча регулиране на машината за пълнене.

13.8. От учебен отдел в Университета „Образование за всеки“ е направена случайна извадка от 25 първокурсници и е получено, че средния им успех от първия семестър е 5,10 с дисперсия 0,059. Знае се, че успехът е нормално разпределена случайна величина.
Като използвате ниво на значимост 0,05, тествайте хипотезата, че успехът на студентите в този университет е над 5,00.

Решение: Интерпретация на данните:

5= константата от хипотезата

5.10=$\overline{x}$ извадковото средно (статистика

0.059= ивадкова дисперсия s^2 (статистика)

25=n=обем на извадката

ниво да значимост 0.05=α

успехът е над 5,00 => едностранен тест, дясностранен

H_0: μ=5 H_1: μ**>5**

НЕизвестна популационна дисерсия/стандартно отклонение и малък обем

Използваме t-разпределението и t- таблицата

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

t(24)=$\frac{5.1-5.00}{\sqrt{0.059}}\sqrt{25} = 2.05847$

При ниво на значимост 0.05 се търси 1-α=0.95 и КО е $(1.71, \infty) \Rightarrow$ Извод

2.05847 е вътре в КО $\Rightarrow$ Отхвърляме H_0. Т.е. успехът е над 5.00

13.10. Агенция по здравословно хранене твърди, че дневната консумация на натрий е повече от 2750 mg. Знае се, че разпределението на дневната консумация на натрий е нормално разпределено. Избрани са по случаен начин 24 жени и е получено, че те приемат средно 2919 mg натрий дневно със стандартно отклонение 1359 mg. При ниво на значимост 0,01 може ли да се отхвърли твърдението на агенцията?

Решение: Интерпретация а данните:

1359= извадково стандартно отклонение s (статистика)

2750= константата от хипотезата

2919=$\bar{x}$ извадковото средно (статистика)

24=n=обем на извадката

ниво да значимост 0.01=α

консумация повече от 2750 mg => едностранен тест, дясностранен

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

H_0: μ=2750 H_1: μ**>2750**

Неизвестна популационна дисперсия и малък обем. Използваме статистиката

t(23)=$\frac{2919-2750}{1359}\sqrt{24} = 0.6092183$

При ниво на значимост 0.01 се търси 1-α=0.99 и КО $(2.5, \infty) \Rightarrow$

Извод 0.609218 НЕ е в КО $\Rightarrow$ Няма основание да отхвърлим H_0. Няма статистическо основание да се смята, че дневната консумация на натрий е повече от 2750 mg.

13.11.Министерството на образованието решава да тества дали учениците със завършено начално образование могат да четат средно поне 150 думи в минута със стандартно отклонение от 15 думи. Избрани са по случаен начин 200 ученици, завършващи средното си образование и на всеки ученик е даден да прочете един и същ текст, като е измерено времето. Намерено е, че средно тези ученици четат по 156 думи в минута, а стандартното отклонение на извадката е 18 думи. Тази информация дава ли статистическо основание да се отхвърли твърдението при предположение за нормалност на изследваната популация?

Решение: Интерпретация а данните:

15 = популационно стандартно отклонение σ(параметър)

150= константата от хипотезата

156=$\bar{x}$ извадковото средно (статистика)

200=n=обем на извадката

Избираме ниво да значимост 0.01=α

консумация повече от 2750 mg => едностранен тест, дясностранен

H_0: μ=150 H_1: μ**>150**

Известна популационна дисперсия. Използваме статистиката

z=$\frac{156-150}{15}\sqrt{200} = 5.65685$

При ниво на значимост 0.01 се търси 1-α=0.99 и КО е $(2.33, \infty) \Rightarrow$

Извод: 5.65685 е в КО ⇒ Отхвърляме H_0. Учениците със завършено начално образование могат да четат средно поне 150 думи в минута