

Toetsen

Vraag 1 (Normale moeilijkheidsgraad)

Open het bestand opgave "*gamerps.csv*". Schaar steen papier is een spelletje waarbij je zelf een keuze maakt tussen deze drie. Is er echter een voorkeur bij onze scoutsleden van de selectietest? De verwachte verdeling is evenredig. Ga na of onze telling afwijkt van de verwachte verdeling

- Welke test ga je hiervoor gebruiken?
- Hoe luidt H_0 ? "De telling in onze steekproef is wel/niet evenredig verdeeld."
- Wat is de waarde van χ^2 ? (3 beduidende cijfers)
- Wat is de kans dat een steekproef een hogere waarde dan χ^2 zou hebben (p-waarde)? (3 beduidende cijfers)
- Kan H_0 met een betrouwbaarheid van 95% verworpen worden?
- Kunnen we dus beweren dat de scoutsleden een voorkeur hebben?

Vraag 2 (Normale moeilijkheidsgraad)

Open het bestand "*hair colorBrussels*". Het gaat om de gegevens over de kleur van het haar van een steekproef van mensen in Brussel. De verwachte verdeling van haarkleuren in Europa zijn:

30% blond, 12% rood, 30% bruin, 25% donker, en 3% zwart.

Ga na of onze telling afwijkt van de verwachte verdeling

- Welke test ga je hiervoor gebruiken?
- Hoe luidt H_0 ? "De haarkleurentelling in onze steekproef wijkt wel/niet af van de
a. verwachte verdeling."
- Welke χ^2 -waarde vind je?
- Welke p-waarde vind je?
- Kan H_0 met een betrouwbaarheid van 95% kan H_0 wel of niet verworpen worden?
- Is er in Brussel een afwijking tov de verwachte verdeling?

Vraag 3 (Normale moeilijkheidsgraad)

We willen weten hoeveel requests een server gemiddeld per dag moet verwerken. Daarvoor doen we een meting. We meten gedurende 30 dagen en tellen iedere dag het aantal requests. We komen op een gemiddelde van 975 en een standaardafwijking gelijk aan 100.

- Tussen welke grenzen ligt dan het gemiddeld aantal requests als we 95% zeker willen zijn?
- Stel dat we hetzelfde gemiddelde en standaardafwijking vonden, maar met een steekproef van 100 dagen. Dan zijn we voor 95% zeker dat het gemiddelde tussen ... en ... ligt.
- Stel dat iemand beweert dat de server 1000 requests per dag moet verwerken. Kan je deze stelling dan ondersteunen of verwerpen in de twee gevallen? Je wil weer een zekerheid van 95%.

Vraag 4 (Normale moeilijkheidsgraad)

Een fabriek maakt zakken cornflakes van 12 kg. Om dit te testen, worden regelmatig steekproeven gedaan. Daarbij worden 100 zakken gewogen. De eerste steekproef levert een gemiddelde van 11,9 kg en een standaardafwijking van 1 kg. We gebruiken een $\alpha=0,01$.

- Welke factor (Z-waarde) zal je gebruiken om het interval op te stellen?
- Moeten we de machine bijstellen?

Een tweede steekproef levert ook een gemiddelde van 11,9 kg, maar een standaardafwijking van 0,1 kg.

- c. Welke factor (Z-waarde) gebruiken we deze keer?
- d. Moeten we de machine deze keer bijstellen?

Vraag 5 (Gevorderde moeilijkheidsgraad)

In opdracht van een kaasfabriek onderzoeken we of sommige leveranciers knoeien met hun melk door er water toe te voegen. We nemen 5 opeenvolgende zendingen melk en kijken op welke temperatuur deze bevriest. We weten dat het vriespunt van melk gelijk is aan $-0,545^{\circ}\text{C}$ met een standaardafwijking van $0,008^{\circ}\text{C}$. Het vriespunt van water is natuurlijk 0°C .

In onze steekproef vinden we een gemiddeld vriespunt van $-0,539^{\circ}\text{C}$. Stel $\alpha = 0,1$

- a. Welke toets gaan we gebruiken?
- b. Moeten we deze toets eenzijdig of tweezijdig toepassen?
- c. Is er met de melk geknoeid?
- d. Wat is de kans dat het vorige antwoord verkeerd is?

Vraag 6 (Normale moeilijkheidsgraad)

We wensen een programmeur aan te werven. We onderwerpen de kandidaten aan een test. We weten dat een goede programmeur een score van gemiddeld 100 haalt op deze test. Enkele (16) studenten van KdG bieden zich aan. We komen op een gemiddelde van 107,3 en een standaardafwijking van 8,0.

Neem $\alpha=0,05$. We vermoeden dat deze groep afwijkt van de gemiddelde populatie.

- a. Welke toets gebruiken we?
- b. Welke factor gaan we gebruiken?

We kunnen een interval opstellen waarbinnen we voor 95% zeker zijn dat KdG studenten gemiddeld scoren.

- c. Wat is de ondergrens van het interval?
- d. Wat is de bovengrens van het interval?
- e. Kunnen we op basis van dit interval stellen dat KdG studenten beter scoren dan gemiddeld?

Vraag 7 (Normale moeilijkheidsgraad)

We wensen een onderzoek uit te voeren naar de favoriete browser van studenten.

We verwachten de volgende verdeling:

- Internet explorer: 8
- Opera: 10
- Mozilla Firefox: 10
- Google chrome: 12

Bij navraag bij 1 klasgroep blijkt de verdeling als volgt:

- Internet explorer: 17
- Opera: 10
- Mozilla Firefox: 8
- Google chrome: 5

- a. Welke test ga je toepassen?
- b. Is deze afwijking binnen onze verwachtingen ($\alpha=0,01$)?
- c. Wat is de kritieke waarde bij deze gegevens?
- d. Welke χ^2 -waarde vind je?
- e. Welke p-waarde vind je?

Vraag 8 (Fun)

Een programmeur heeft een Python klasse geschreven om random gehele getallen tussen 0 en 10 te genereren. Deze code ziet er als volgt uit:

```
import math

class Random:
    def __init__(self)-> None:
        self.i = 0
        self.j = 0

    def random_between_0_and_10(self) -> int:
        r = abs(math.sin(51 * self.i ** 2 + math.cos(80 * self.j)) +
                math.sin(300 * self.j + 3 * math.sin(111 * self.i - self.j)))
        self.j += 1
        if self.j > 100:
            self.i +=1
            if self.i > 100:
                self.i = 0
            self.j = 0

        return math.floor(r * 5.5)
```

Schrijf een Python programma dat 1100 random getallen genereert en ga na of deze random getallen generator goed werkt (we verwachten dat ieder getal 0-10 evenveel voorkomt). Gebruik $\alpha=0.05$.

1. Welke χ^2 -waarde vind je?
2. Welke p-waarde vind je?
3. Wat is het betrouwbaarheidsinterval voor $\alpha=0.05$?
4. Welk besluit kan je trekken?