

VU-stat opgaven 8 juni 2016

▣ Opgave 1

Open het bestand 'Voetlengtes' in VU-stat, je vindt daarin de voetlengtes in centimeter van honderd mannen en honderd vrouwen.

- a) Bepaal van elk van beide groepen zowel de gemiddelde voetlengte als de standaarddeviatie van de voetlengtes, beide in 1 decimaal nauwkeurig.
- b) Maak bij elke van beide groepen een histogram en laat aan de hand daarvan zien dat de voetlengtes bij goede benadering normaal zijn verdeeld.

Maak daartoe eerst frequentietabellen van alle voetlengtes afgerond op gehele centimeters.

Werk nu verder met normale verdelingen als model voor de voetlengtes.

- c) Hoeveel procent van de mannen heeft grotere voeten dan de helft van de vrouwen?
- d) Wat is de kleinste voetlengte die voorkomt in de groep van de 2,5 procent vrouwen met de grootste voeten?
- e) Wat is de grootste voetlengte die voorkomt in de groep van de 2,5 procent mannen met de kleinste voeten?

▣ Opgave 2

Open het bestand Oldfaithful2.

De Oldfaithful is een geiser in Yellowstone park in de VS. Ze hebben een eigen website.

Er zijn in oldfaithful2.vus twee variabelen: lengte van de erupties van de geiser en de wachttijd tot de volgende eruptie.

- a. Maak een dotplot van eruptielengte en maak een dotplot van de wachttijd. Je ziet dat beide dotplots een bergje geven. Wat voor vermoeden zou je hier kunnen formuleren?

Dit vermoeden willen we onderzoeken; er kan samenhang zijn maar teoval is niet uit te sluiten.

- b. Maak eerst een puntenwolk, dit leidt tot hetzelfde vermoeden.

We hebben een nieuwe variabele nodig: een korte of een lange cyclus.

- c. definieer een nieuwe variabele (via Data, variabele) en vul deze met de waarde 1 voor een korte eruptieduur en met de waarde 2 voor een lange eruptieduur (Data, hercoderen)

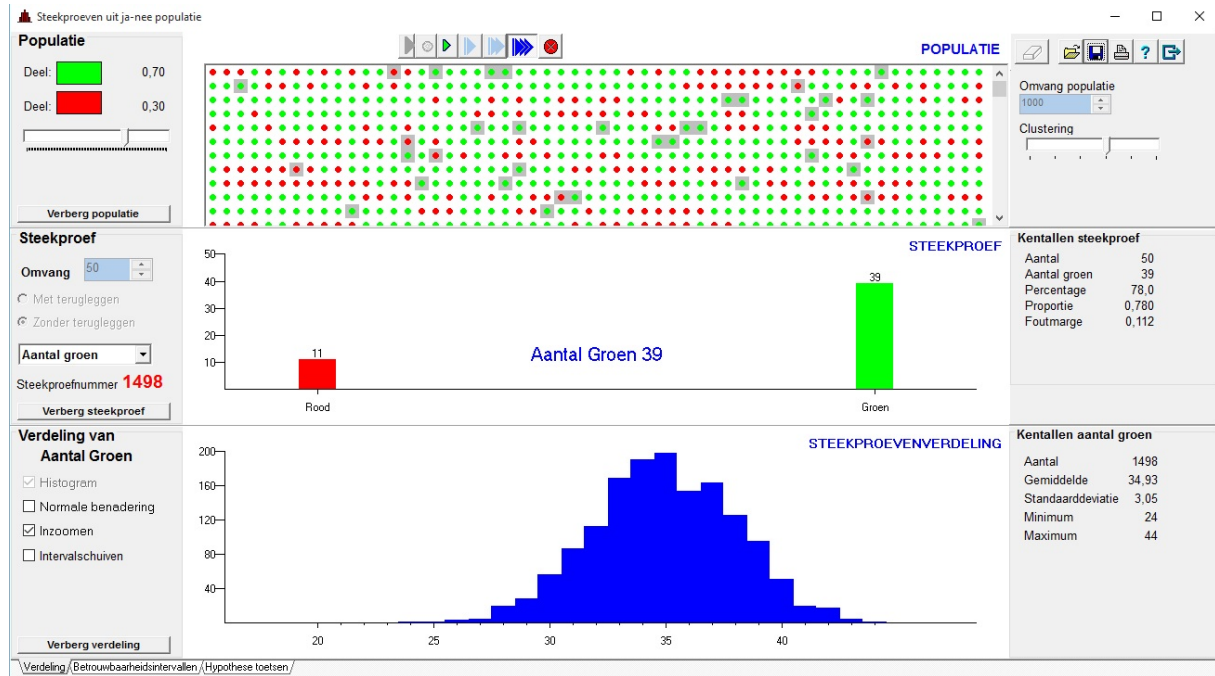
- d. Nu kun je de dataset splitsen op lange en korte eruptieduur en weer de puntenwolk bekijken.

- e. Welke kwalitatieve conclusie kun je nu trekken?

Opgave 3

Ga naar VuStat → steekproeven → steekproeven uit ja-nee populatie.
Stel de proportie op 0,70; de populatieomvang op 10000; de omvang van de steekproef op 50.

Laat dit 1000 keer naspelen (je maakt 1000 simulaties); je krijgt zoiets:



Voer dit nog enkele malen (bijvoorbeeld 5 keer) uit waarbij je steeds 1000 simulaties laat uitvoeren. Kijk ook naar andere geheime proporties dan 0,70. Trek conclusies over de vorm van de grafieken die je voor de steekproefproporties krijgt (zie rechtsonder in bovenstaand plaatje).

Opgave 4

We gaan nu werken met het bestand Beroepsbevolking. We kijken naar het verschil tussen mannen en vrouwen die werkzaam zijn in de medische wereld. De bedoeling is dat je drie uitspraken controleert:

Er is een groot verschil tussen mannen en vrouwen voor wat betreft:

1. arbeidsduur;
2. onderwijsniveau;
3. proportie die meer zou willen werken.

a). Controleer deze uitspraken.

Het gegevensbestand bevat variabelen die een samenhang zouden kunnen vertonen.

Een logische gedachte lijkt dat juist mensen met een fulltime baan minder zouden willen werken en dat juist mensen met een parttime baan meer zouden willen werken.

b). Onderzoek met behulp van een kruistabel of de variabelen *arbeidsduur* en *meer of minder willen werken* op deze manier samenhangen.

Opgave 5

Ga in VU-stat naar 'veel steekproeven trekken'. Gebruik het bestand Temperaturoverzet. Hierin zie je gegevens over het weer vanaf 1874.

- Gebruik als variabele 'annual' (je mag een andere kolom maken met de temperatuur in °C, omrekenen via: $^{\circ}\text{C} = (\text{F}-32)/1,8$)
- gebruik als uitvoer 'gemiddelde'
- Wat is de populatie?
- Neem eerst een aantal keren een kleine steekproef uit de populatie.

Kies eerst voor steekproeven van drie verschillende jaren. Laat 100 keer een dergelijke steekproef trekken.

Verwerk de resultaten in een staafdiagram. Ontstaat er een klokvormig staafdiagram?

Laat vervolgens steekproeven trekken van 30 verschillende jaren uit de populatie en bekijk weer het staafdiagram van het steekproefgemiddelde. Ontstaat een klokvormig staafdiagram?

Bekijk ook nog eens het staafdiagram van de hele populatie (alle gegeven jaren).

Trek conclusies. Komen die overeen met de conclusie hieronder?

Conclusie

Voor kleine steekproeven ontstaat geen klokvormig staafdiagram van de steekproefgemiddelden, maar voor grote steekproeven is dat bij benadering wel het geval.

Onderzoek nu een andere zelfgekozen variabele. Controleer of ook voor deze andere variabele geldt dat bij een grote steekproefomvang het staafdiagram van het steekproefgemiddelde bij benadering klokvormig is (ongeacht de verdeling van de populatie).

Opgave 6

Neem een van de datasets die tot je beschikking staan en bedenk zelf een opdracht voor leerlingen bij deze dataset.

Of

Formuleer een onderzoek dat je in de klas zou kunnen uitvoeren en bedenk wat leerlingen kunnen onderzoeken.