STATISTIEK voor bedrijfskundigen I
Oefeningenbundel

Bachelor of Science in de handelswetenschappen
Universiteit Gent

Academiejaar 2019-2020

Inhoudsopgave

Opgaven

Hoofdstuk 6

Verdelingen van steekproefgrootheden

6.1 Basisconcepten

Oefening 6.1. De kansverdeling hieronder beschrijft een populatie van metingen die de waarden $x = 0, 2, 4, 6$ kunnen aannemen met bijbehorende relatieve frequenties $p(x)$:

x	0	2	4	6
$p(x)$	0,2	0,1	0,5	0,2

a. Beschouw alle aselecte steekproeven van omvang $n = 2$ en bepaal de verdeling van het bijbehorende steekproefgemiddelde $\bar{x}$.

b. Teken het kanshistogram van de steekproefverdeling van $\bar{x}$.

c. Wat is de kans dat $\bar{x}$ groter is dan 3?

d. Bepaal het populatiegemiddelde μ.

e. Bepaal de verwachtingswaarde van $\bar{x}$. Laat zien dat $\bar{x}$ een zuivere schatter is van μ.

Oefening 6.2. Beschouw een populatie gekenmerkt door de volgende kansverdeling

x	0	1
$p(x)$	$\frac{1}{2}$	$\frac{1}{2}$

a. Bepaal de steekproefverdeling van het steekproefgemiddelde voor een aselecte steekproef van $n = 3$ waarnemingen uit deze populatie.

b. Bepaal de steekproefverdeling van de mediaan van een steekproef van $n = 3$ waarnemingen uit deze populatie.

c. Bepaal het populatiegemiddelde μ en laat zien dat zowel het steekproefgemiddelde als de steekproefmediaan een zuivere schatter van μ is.

d. Bereken de varianties van het steekproefgemiddelde en de steekproefmediaan.

e. Welke schatter zou je kiezen om μ te schatten? Leg uit.

Oefening 6.3. Een aselecte steekproef van 64 waarnemingen wordt genomen uit een populatie met een gemiddelde gelijk aan 20 en een standaardafwijking gelijk aan 16.

a. Geef het gemiddelde en de standaardafwijking van de steekproefverdeling van $\bar{x}$.

b. Beschrijf de vorm van de steekproefverdeling van $\bar{x}$. Hangt je antwoord af van de steekproefomvang?

c. Bepaal de kans dat $\bar{x}$ kleiner is dan 16.

d. Bepaal de kans dat $\bar{x}$ gelegen is tussen 18 en 22.

e. Voor welke waarde a geldt dat $P(\bar{x} > a) = 0{,}10$?

Oefening 6.4. Een omvangrijke aselecte steekproef van 900 observaties wordt genomen uit een populatie met gemiddelde en variantie gelijk aan 100. Tussen welke twee waarden zal het steekproefgemiddelde met 95% kans gelegen zijn?

Oefening 6.5. Een aselecte steekproef van 100 waarnemingen wordt genomen uit een populatie met gemiddelde 30 en standaardafwijking 16. Benader de volgende kansen:

a. $P(\bar{x} \geq 28)$

b. $P(25{,}2 \leq \bar{x} \leq 26{,}8)$

c. $P(\bar{x} < 32{,}4)$

6.2 Toepassingen

Oefening 6.6. Aan het einde van de twintigste eeuw was de kans van werknemers om een lange tijd bij dezelfde werkgever te blijven veel kleiner dan die van hun ouders in de vorige generatie (Georgia Trend, december 1999). Realiseren studenten van vandaag zich dat de werkplek die ze binnenkort zullen betreden sterk verschilt van die van hun ouders? Om deze vraag te beantwoorden hebben onderzoekers van het Terry College of Business van de Universiteit van Georgia een steekproef genomen van 344 studenten handelswetenschappen en hebben hen de volgende vraag gesteld: wat is het maximum aantal jaar dat je bij één en dezelfde werkgever denkt te blijven werken? Deze steekproef gaf een gemiddelde van $\bar{x} = 19{,}1$ jaar en $s = 6$ jaar. Stel dat de steekproef van studenten aselect is gekozen uit de 5800 studenten van het Terry College.

a. Beschrijf de steekproefverdeling van $\bar{x}$.

b. Als het populatiegemiddelde 18,5 jaar is, wat is dan de kans dat het steekproefgemiddelde groter is dan 19,1 jaar?

c. Zelfde vraag als (b), maar voor een populatiegemiddelde van 19,5 jaar.

d. Als $P(\bar{x} \geq 19{,}1) = 0{,}5$, wat is dan het populatiegemiddelde?

e. Als $P(\bar{x} \geq 19, 1) = 0, 2$, is het populatiegemiddelde dan groter of kleiner dan 19,1 jaar? Licht je antwoord toe.

Oefening 6.7. Een winkelier die wil weten wanneer hij een order moet plaatsen om de voorraad van een product die uitgeput raakt aan te vullen, moet rekening houden met de doorlooptijden voor de producten. De *doorlooptijd* (Engels: *lead time*) is de tijd die verloopt tussen het plaatsen van een order en het beschikken over het product zodat aan de vraag van de klant kan worden voldaan. Dit omvat de tijd nodig voor het plaatsen van de order, het ontvangen van de zending van de leverancier, het controleren van de ontvangen artikelen en het opnemen ervan in de voorraad (Clauss, Applied Management Science and Spreadsheet modeling, 1966).
De inkoopafdeling van een grote warenhuisketen is geïnteresseerd in de gemiddelde doorlooptijd μ voor een bepaalde leverancier van herenkleding, en neemt daarom een aselecte steekproef van 50 doorlooptijden voor deze leverancier en vindt daarvoor een waarde $\bar{x} = 43$ dagen.

a. Beschrijf de vorm van de steekproefverdeling van $\bar{x}$.

b. Als $\mu = 40$ en $\sigma^2 = 112{,}5$, hoe groot is dan de kans dat een tweede aselecte steekproef van 50 doorlooptijden een waarde voor $\bar{x}$ zou opleveren die groter is dan of gelijk is aan 43?

c. Als je de waarden voor μ en σ uit (b) gebruikt, hoe groot is dan de kans dat een steekproef van 50 doorlooptijden een steekproefgemiddelde heeft dat binnen het interval $[39, 41]$ ligt?

Oefening 6.8. Een belegger wil eenzelfde bedrag investeren in elk van n verschillende aandelen. Het jaarlijks rendement van elk aandeel heeft een gemiddelde $\mu = 10\%$ en een standaardafwijking $\sigma = 4\%$. Het jaarlijks rendement voor de belegger van de portefeuille van n aandelen is $\bar{r} = \frac{1}{n}\sum_{i=1}^{n} r_i$.

a. Welk jaarlijks rendement kan de belegger verwachten? Hangt dit af van n?

b. Bepaal de variantie van het jaarlijks rendement. Dit getal is een maat voor het risico dat de belegger loopt.

c. Neemt het beleggersrisico toe of af met n? Verklaar.

d. Bepaal n zodat de belegger met minstens 95% zekerheid kan voorspellen dat het jaarlijks rendement van zijn aandelenportefeuille tussen 9% en 11% zal liggen.

Oefening 6.9. Een bottelaar van frisdrank heeft flessen nodig met een interne druksterkte van ten minste 150 pounds per square inch (psi). Een groothandelaar van flessen claimt dat zijn productieproces flessen oplevert met een gemiddelde interne druksterkte van 157 psi en een standaardafwijking van 3 psi. Om de bewering van de groothandelaar te toetsen, neemt de bottelaar een aselecte steekproef van 40 flessen uit de 10 000 meest recent geproduceerde flessen, meet de interne druksterkte van elke fles, en vindt een gemiddelde waarde voor de druksterkte die 1,3 psi lager ligt dan het procesgemiddelde dat door de groothandelaar wordt beweerd.

a. In de onderstelling dat de bewering van de groothandelaar juist is, hoe groot is dan de kans dat je een steekproefgemiddelde vindt dat minstens 1,3 psi lager ligt dan het procesgemiddelde? Wat suggereert je antwoord met betrekking tot de geldigheid van de claim van de groothandelaar?

b. Onderstel dat de processtandaardafwijking 3 psi is, zoals door de groothandelaar wordt beweerd, maar het procesgemiddelde is 156 psi. Is het waargenomen resultaat van de steekproef dan meer of minder waarschijnlijk zijn dan in (a)? En als in plaats daarvan de standaarddeviatie 6 psi zou zijn?

Oefening 6.10. De *British Medical Journal* (aug 17, 2002) publiceerde een onderzoek waarbij de effectiviteit van zeep als ontsmettingsmiddel bij het handen wassen werd vergeleken met alcohol. Werknemers in de gezondheidszorg die zeep gebruiken hebben na het handen wassen nog gemiddeld 69 bacteriesoorten op hun handen, met een standaardafwijking van 106 bacteriesoorten. Bij het handen wassen met alcohol blijven gemiddeld nog 35 bacteriesoorten over, met een standaardafwijking van 59. In een aselecte steekproef van 81 werknemers uit de gezondheidszorg die allen hetzelfde ontsmettingsmiddel gebruiken bij het handen wassen, blijkt het gemiddeld aantal overblijvende bacteriesoorten gelijk te zijn aan 49. Werd in deze steekproef zeep of alcohol gebruikt als ontsmettingsmiddel?

Oplossingen

Hoofdstuk 6

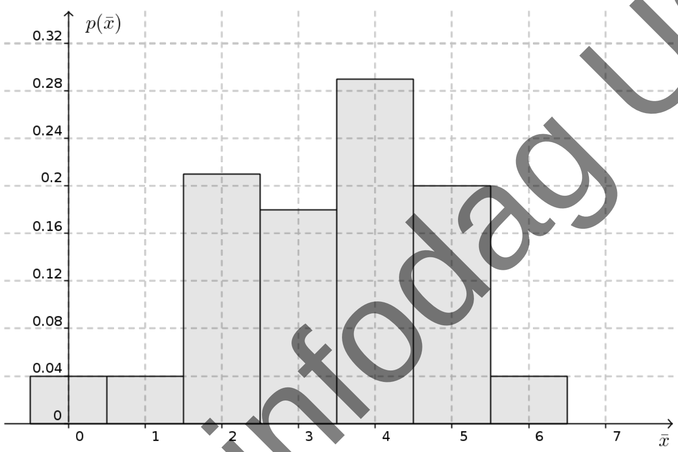
Verdelingen van steekproefgrootheden

Oefening 6.1.

a.

$\bar{x}$	0	1	2	3	4	5	6
$p(\bar{x})$	0,04	0,04	0,21	0,18	0,29	0,20	0,04

b. Kanshistogram:



c. 0,53

d. $\mu = 3{,}4$

e. $E(\bar{x}) = 3{,}4$
$\bar{x}$ is een zuivere schatter van μ, want $E(\bar{x}) = \mu$.

Oefening 6.2.

a.

$\bar{x}$	0	$\frac{1}{3}$	$\frac{2}{3}$	1
$p(\bar{x})$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

b.

m	0	1
$p(m)$	$\frac{1}{2}$	$\frac{1}{2}$

c. $\mu = \frac{1}{2}$
$E(\bar{x}) = \frac{1}{2}$
$E(m) = \frac{1}{2}$

Zowel $\bar{x}$ en m zijn voor deze populatie zuivere schatters van μ omdat hun verwachtingswaarden gelijk zijn aan μ.

d. $\sigma^2_{\bar{x}} = E[(\bar{x} - E(\bar{x}))^2] = \frac{1}{12}$
$\sigma^2_m = E[(m - E(m))^2] = \frac{1}{4}$

e. Traditioneel wordt gekozen voor de schatter die het minst varieert over de steekproeven heen. Kies dus $\bar{x}$, omdat zij de kleinste variantie heeft (zie (d)).

Oefening 6.3.

Gegeven: $n = 64$, $\mu = 20$, $\sigma = 16$.

a. $\mu_{\bar{x}} = \mu = 20$, $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = 2$

b. Wegens de centrale limietstelling is $\bar{x}$ bij benadering normaal verdeeld voor grote waarden van n. Hier is $n = 64$ voldoende groot.

c. $P(\bar{x} < 16) \approx 0{,}0228$

d. $P(18 < \bar{x} < 22) \approx 0{,}6826$

e. $a \approx 22{,}56$

Oefening 6.4.

$\mu_{\bar{x}} - 2\sigma_{\bar{x}} \approx 99{,}33$ en $\mu_{\bar{x}} + 2\sigma_{\bar{x}} \approx 100{,}67$.

Oefening 6.5.

a. 0,8944

b. 0,0215

c. 0,9332

Oefening 6.6.

a. Door de centrale limietstelling heeft $\bar{x}$ bij benadering een normale verdeling met $\mu_{\bar{x}} = \mu$ en $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{344}}$.

b. Gegeven $\mu = 18{,}5$. Omdat σ niet gekend is zullen we dit schatten met $s = 6$.
$P(\bar{x} > 19{,}1) \approx P(z > \frac{19{,}1-18{,}5}{6/\sqrt{344}}) \approx P(z > 1{,}85) = 0{,}5 - 0{,}4678 = 0{,}0322$

c. Gegeven $\mu = 19{,}5$. Omdat σ niet gekend is zullen we dit schatten met $s = 6$.
$P(\bar{x} > 19{,}1) \approx P(z > \frac{19{,}1-19{,}5}{6/\sqrt{344}}) \approx P(z > -1{,}24) = 0{,}5 + 0{,}3925 = 0{,}8925$

d. $\mu = 19{,}1$

e. Kleiner. Als μ groter zou zijn dan 19,1, dan zou $P(\bar{x} \geq \mu) < P(\bar{x} \geq 19{,}1) = 0{,}2$. Dit is onmogelijk, want door de centrale limietstelling is $\bar{x}$ bij benadering normaal verdeeld en zou dus moeten gelden dat $P(\bar{x} \geq \mu) \approx 0{,}5$.

Oefening 6.7.

a. Door de centrale limietstelling heeft $\bar{x}$ bij benadering een normale verdeling met $\mu_{\bar{x}} = \mu$ en $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{50}}$. Ze is dus klokvormig.

b. $P(\bar{x} \geq 43) \approx 0{,}0228$

c. $P(39 \leq \bar{x} \leq 41) \approx 0{,}4972$

Oefening 6.8.

a. $\mu_{\bar{r}} = \mu = 10$. Dit hangt niet af van n.

b. $\sigma^2_{\bar{r}} = \frac{\sigma^2}{n} = \frac{16}{n}$

c. Het beleggersrisico neemt af met n, omwille van de formule in (b).

d. Zoek n zodat $P(9 < \bar{r} < 11) \geq 0{,}95$. Antwoord: $n \geq 62$.

Oefening 6.9.

Gegeven: $n = 40$, $\mu = 157$ en $\sigma = 3$.

a. $P(\bar{x} \leq \mu - 1{,}3) \approx 0{,}0031$
Het vinden van een steekproefgemiddelde dat minstens 1,3 psi onder het procesgemiddelde ligt is een zeldzame gebeurtenis onder de bewering van de groothandelaar. Misschien spreekt deze niet de waarheid?

b.
- $\mu = 156$, $\sigma = 3$
  Als μ verandert, maar σ niet, dan vinden we dezelfde kans als in (a). De kans $P(\bar{x} \leq \mu - 1{,}3)$ is onafhankelijk van μ.
- $\mu = 157$, $\sigma = 6$
  $P(\bar{x} \leq \mu - 1{,}3) \approx P(z \leq \frac{-1{,}3}{6/\sqrt{40}}) \approx P(z \leq -1{,}37) = 0{,}5 - 0{,}4147 = 0{,}0853$
  Deze kans hangt wel af van σ en is groter dan deze in (a). Het gevonden steekproefgemiddelde is dus waarschijnlijker dan in (a).

Oefening 6.10.

Indien met zeep:
$\mu_{\bar{x}} = \mu_{\text{zeep}} = 69$, $\sigma_{\bar{x}} = \frac{\sigma_{\text{zeep}}}{\sqrt{n}} = \frac{106}{\sqrt{81}} \approx 11{,}78$
Het steekproefgemiddelde ($\bar{x} = 49$) is minder dan 2 standaardafwijkingen verwijderd van $\mu_{\bar{x}}$. De kans dat dit gebeurt is ongeveer 95% (empirische regel).

Indien met alcohol:
$\mu_{\bar{x}} = \mu_{\text{alc}} = 35$, $\sigma_{\bar{x}} = \frac{\sigma_{\text{alc}}}{\sqrt{n}} = \frac{59}{\sqrt{81}} \approx 6{,}56$
Het steekproefgemiddelde ($\bar{x} = 49$) is verder dan 2 standaardafwijkingen van $\mu_{\bar{x}}$ verwijderd. De kans dat dit gebeurt is slechts 5% (empirische regel).

Omdat de waargenomen waarde ($\bar{x} = 49$) bij de alcoholhypothese een zeldzame gebeurtenis is (en niet bij de zeephypothese), zou je kunnen vermoeden dat de steekproef waarschijnlijk zeep heeft gebruikt als ontsmettingsmiddel.