**Praktische statistiek – uitwerkingen hfdst. 5**

5.1a

Rekenkundig gemiddelde = $\frac{50 × 17 + 100 × 22 +…+ 500 × 83}{22 + 17 +…+ 83}=\frac{77.500}{250}=$ 310 euro

Modus = 500 euro (meest voorkomende waarneming)

Mediaan = $\frac{125ste waarneming + 126ste waarneming}{2}=\frac{300+300}{2}= $300 euro

5.1b

Variantie = $\frac{(50-310)^{2} × 17 + (1000-310)^{2} × 22+…+ (500-310)^{2} × 83}{250}=$ 25.980 euro2

s.d. = $\sqrt{variantie}= √25.980$ $≈$ 161,2 euro

5.1c

In de berekening van zowel het r.g. als de variantie geldt: teller en noemer worden beide verdubbeld, dus uitkomsten veranderen niet. Dan zal dus ook de uitkomst van de s.d. niet veranderen.

5.2a

Gemiddelde brandduur = $\frac{4.400 ×1 +4.600 × 5+…+5.600 ×0}{1 +5 +…+0}=\frac{500.000}{100}=$ 5.000 uur

Variantie = $\frac{(4.400 – 5.000)^{2} ×1 + (4.600 - 5.000)^{2} ×5 +…+ (5.600 - 5.000)^{2} × 0}{100}=$ 42.400 uur2

s.d. = $√42.400$ = 205,9 dus in honderden nauwkeurig: 200 uur

5.2b

Variatiecoëfficiënt = $\frac{s.d.}{r.g.}$ = $\frac{200}{5.000}$ × 100% = 4% (dimensieloze grootheid)

5.2c

|  |  |
| --- | --- |
| Brandduur (uren) | Cum. Rel. Freq.  |
| 4.300 -< 4.500 |  1% |
|  -< 4.700 |  6% |
|  -< 4.900 |  31% |
|  -< 5.100 |  69% |
|  -< 5.300 |  93% |
|  -< 5.500 | 100% |
|  -<5.700 | 100% |

Tot 5.000: 31 + $\frac{69 -31}{2}=$ 50% en tot 5.200: 69 + $\frac{93 -69}{2}=$ 81%;

Dus 31% ligt tussen 5.000 en 5.200 branduren.

5.2d

|  |  |
| --- | --- |
| Brandduur (uren) | Cum. Rel. Freq. van hoog naar laag |
| > 4.300 | 100% |
| > 4.500 |  99% |
| > 4.700 |  94% |
| > 4.900 |  69% |
| > 5.100 |  31% |
| > 5.300 |  7% |
| > 5.500 en <5.700 |  0% |

r.g. + 2 × s.d. = 5.000 + 2 × 200 = 5.400 branduren

> 5.300 uren: 7% en > 5.500 uren: 0%

Dus meer dan 5.400 uren naar (grove) schatting 3,5%, dus kans = 3,5%.

Evenzo r.g. − 2 × s.d. = 5.000 − 2 × 200 = 4.600 branduren

> 4.500 uren: 99% en > 4.700 uren: 94%

Dus meer dan 4.600 uren grof geschat 96,5% ofwel 3,5% minder dan 4.600 uren;

in totaal wijkt circa 7% verder van het r.g. af dan tweemaal de s.d.

5.2e

Hier moet dezelfde berekening worden uitgevoerd als in 5.2d. Antwoord dus 3,5%.

5.2f

Zie tabel bij 5.2d

r.g.+ s.d. = 5.200 en r.g.– s.d. = 4.800

> 4.700 uren: 94% en > 4.900 uren: 69%; dus > 4.800 uren: 81,5%

> 5.100 uren: 31% en > 5.500 uren: 7%; dus > 5.200 uren: 19%

Dus tussen 4.800 en 5.200 uren: 81,5 – 19 = 62,5%

5.2g

Zie tabel bij 5.2d

$\frac{? - 5.100}{5.300 - 5.100}$ $= \frac{15,9 -31}{7-31}$ dus ? = 5.100 + $\frac{15,9 -31}{7-31}$ × 200 = 5.226 branduren

5.2h

Bij benadering symmetrisch;

hier wijkt 62% van de uitkomsten niet meer dan eenmaal de s.d. af van het r.g. (bij een normale verdeling is dit circa 68%)

rechteroverschrijdingskans bij tweemaal de s.d. is hier 3,5%. (bij een normale verdeling is dit circa 2,5%)

5.3a

De kans dat de omzet groter is dan 170.000 euro rekenen we terug naar de standaard: de afwijking bedraagt 20.000 bij een s.d. van 12.000, dit is dus 20/12 ofwel circa 1,67 keer de s.d. De rechteroverschrijdingskans is de gevraagde kans en bedraagt circa 5%.

5.3b

2σ – interval omzet is <126.000 ; 174.000>

2σ – interval kosten is <92.000 ; 128.000>

in beide gevallen ligt ca. 95% van de uitkomsten in het betreffende interval; daarnaast is de “overlap” slechts 2.000 euro.

5.3c

Bij een linkeroverschrijdingskans van 10% hoort een afwijking van circa 1,28 keer de s.d. (gebruik gemaakt van de inverse normale verdeling: NORM.S.INV(0,1) = −1,28) dus: te noemen omzet = gemiddelde – 1,28 × s.d. = 150.000 – 1,28 × 12.000 en dit komt uit op € 134.640.

N.B.

Deze opgave valt eigenlijk niet binnen de leerstof uit het boek (excuses!); er kan een mouw aangepast worden als bv. in Excel of met de GR de inverse normale functie wordt gebruikt (NORM.S.INV bij Excel, daarbij uitgaande van μ = 0 en s.d. = 1) met als argument 0,1: het antwoord is dan circa -1,28.

M.b.v. ‘trial and error’ kan het ook, maar is uiteraard arbeidsintensiever; bv.
NORM.S.VERD(-1) geeft circa 0,1587 (= 15,87%), NORM.S.VERD(-1,5) geeft
circa 0,0668 (= 6,68%) enz.

Dan kun je blijven zoeken totdat je zo dicht als wenselijk is bij -1,28 uitkomt, dus:

|  |  |
| --- | --- |
| z-waarde | Linker overschrijdingskans |
| -1  | 0,158655254 |
| -1,5 | 0,066807201 |
| -1,3    | 0,096800485 |
| ……..  | 0,10… |

Ook kan bij een aantal veelvoorkomende overschrijdingskansen de bijbehorende z-waarde uit het hoofd worden geleerd dan wel als gegeven bij de opgave worden gevoegd. Dus:

|  |  |
| --- | --- |
| Linker overschrijdingskans in procenten | z-waarde (drie decimalen) |
| 1%  | −2,326 |
| 2,5% | −1,960 |
|   5% | −1,645 |
| 10%  | −1,282 |
| 90% | 1,282 |
| 95% | 1,645 |
| 97,5% | 1,960 |
| 99% | 2,326 |

Merk op dat de laatste vier z-waarden dezelfde zijn als wanneer gesproken zou worden van rechteroverschrijdingskansen van respectievelijk 10%, 5%, 2,5% en 1%.

5.4a

We moeten kijken naar de linkeroverschrijdingskans die hoort bij een afwijking van eenmaal de s.d. deze kans is bij benadering gelijk aan 15,9%.

5.4b

ondergrens = r.g. – 2σ = 7,5 - 2×7,5 = -7,5%

5.5a

gemiddelde = $\frac{1.100+1.700}{2}=1.400$

5.5b

6σ = 1.700 – 1.100 = 600 dus σ = 100

5.5c

r.g. - 2σ = 1.200 en r.g. + 2σ = 1.600, dus 2σ-interval is <−1.200 ; 1.600>