Voortraject Statistiek

Prof. dr. Ellen Vandervieren

ellen.vandervieren@uantwerpen.be
CST, Venusstraat 35, bureau 107

Universiteit Antwerpen

Voortraject - Vrijdag

- Faculteiten
- Frequentietabel - Gemiddelde - Standaardafwijking - Grafische voorstelling
- Empirische kansdefinitie o.b.v. limiet
- Kans o.b.v. bepaalde integraal

Faculteiten

- $3! = 3 \cdot 2 \cdot 1$ ('3 faculteit')
- $4! = 4 \cdot 3 \cdot 2 \cdot 1$ ('4 faculteit')
- $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$
- $\ldots$
- $100! = 100 \cdot 99 \cdot 98 \cdot 97 \cdot 96 \cdot \ldots \cdot 2 \cdot 1$
- **Algemeen:** $n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 2 \cdot 1$
- **Speciaal geval:** $0! = 1$

Faculteiten

Uitdrukkingen met faculteiten vereenvoudigen:

- vb. $\frac{4!}{3!} = \frac{4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1} = \frac{4 \cdot \cancel{3} \cdot \cancel{2} \cdot 1}{\cancel{3} \cdot \cancel{2} \cdot 1} = 4$

  OF: $\frac{4!}{3!} = \frac{4 \cdot 3!}{3!} = \frac{4 \cdot \cancel{3!}}{\cancel{3!}} = 4$
- vb. $\frac{10!}{3! \cdot 7!} = \frac{10 \cdot 9 \cdot 8 \cdot \ldots \cdot 1}{3 \cdot 2 \cdot 1 \cdot 7 \cdot 6 \cdot 5 \cdot \ldots \cdot 2 \cdot 1} = \ldots$

  OF: $\frac{10!}{3! \cdot 7!} = \frac{10 \cdot 9 \cdot 8 \cdot 7!}{3! \cdot 7!} = \frac{10 \cdot 9 \cdot 8 \cdot \cancel{7!}}{3! \cdot \cancel{7!}}$

  $= \frac{10 \cdot 9 \cdot 8}{3 \cdot 2 \cdot 1} = 5 \cdot 3 \cdot 8 = 120$

Faculteiten

Uitdrukkingen met faculteiten anders schrijven:

- $4! = 4 \cdot (3 \cdot 2 \cdot 1) = 4 \cdot 3!$
- $5! = 5 \cdot (4 \cdot 3 \cdot 2 \cdot 1) = 5 \cdot 4!$
- $10! = 10 \cdot (9 \cdot 8 \cdot \ldots \cdot 2 \cdot 1) = 10 \cdot 9!$
- **Algemeen:** $n! = n \cdot (n-1)!$

Faculteiten

Uitdrukkingen met faculteiten anders schrijven:

- $5! = 5 \cdot 4 \cdot (3 \cdot 2 \cdot 1) = 5 \cdot 4 \cdot 3!$
- $10! = 10 \cdot 9 \cdot 8 \cdot (7 \cdot 6 \cdot \ldots \cdot 2 \cdot 1) = 10 \cdot 9 \cdot 8 \cdot 7!$
- $15! = 15 \cdot 14 \cdot 13 \cdot 12 \cdot (11 \cdot 10 \cdot \ldots \cdot 2 \cdot 1)$
  $= 15 \cdot 14 \cdot 13 \cdot 12 \cdot 11!$
- **Algemeen:**
  $n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots (n-p) \cdot (n-p-1)!$
  OF:
  $n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots (n-k+1) \cdot (n-k)!$

Faculteiten

Oefenmoment:

vb. $\dfrac{20!\,3!}{6!\,17!}$

vb. $\dfrac{12!\,0!}{8!\,5!}$

vb. $\dfrac{4!\,7!}{10!\,3!}$

Toepassing in Statistiek

Combinatie:

- vb. 5 vrienden (Ahmed, Bart, Cleo, Diane, Eli) willen in commissie zetelen
- Er zijn slechts 2 stoelen beschikbaar
- Op hoeveel manieren kan de afvaardiging van 2 personen worden samengesteld?
- We moeten 2 personen kiezen uit een verzameling van 5 personen (A, B, C, D, E) waarbij
  - er geen teruglegging is (een persoon wordt hoogstens één maal gekozen)
  - de volgorde van geen belang is (vb. combinatie (A,C) is zelfde als (C,A))

Toepassing in Statistiek

Combinatie:

- Mogelijke combinaties?
- (A,B) (A,C) (A,D) (A,E) (B,C)
  (B,D) (B,E) (C,D) (C,E) (D,E)
- Aantal combinaties van 2 personen uit verzameling van 5 personen = 10
- Achterliggende formule?
  $\frac{5!}{2!\cdot(5-2)!} = \frac{5\cdot4\cdot3!}{2!\cdot3!} = \frac{5\cdot4}{2} = 10$
- **Notatie:** $C_5^2 = \binom{5}{2} = \frac{5!}{2!\cdot(5-2)!} = 10$

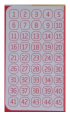
Toepassing in Statistiek

Combinatie:

- Wanneer k elementen worden gekozen uit een verzameling van n elementen waarbij
  - er geen teruglegging is (ieder element wordt hoogstens één maal gekozen)
  - de volgorde van geen belang is
- **Notatie:** C_n^k = aantal combinaties van k elementen uit verzameling van n elementen
- **Formule:** $C_n^k = \binom{n}{k} = \frac{n!}{k!\cdot(n-k)!}$
- $\binom{n}{k}$ is een binomiaalcoëfficiënt ('n over k')

Toepassing in Statistiek

- vb. Lotto biljet:

- 'Enkelvoudig lotto biljet': 

Toepassing in Statistiek

- vb. enkelvoudig Lotto biljet: op hoeveel manieren kan je 6 uit 45 getallen aankruisen?
  - geen teruglegging: een getal kan niet 2x gekozen worden
  - volgorde waarin getallen worden aangekruist heeft geen belang
- aantal combinaties berekenen van 6 uit 45
- $C_{45}^6 = \binom{45}{6} = \frac{45!}{6!\cdot(45-6)!} = \frac{45!}{6!\cdot39!}$

$$= \frac{45\cdot44\cdot43\cdot42\cdot41\cdot40\cdot39!}{6!\cdot39!} = \frac{45\cdot44\cdot43\cdot42\cdot41\cdot40}{6!}$$

$= 8\,145\,060$

- kans op 6 juiste cijfers is $\frac{1}{8\,145\,060}$

Toepassing in Statistiek

- vb. op hoeveel manieren kan je uit een groep van 12 vrienden een quizploeg samenstellen van 5 personen?
  - geen teruglegging: een persoon kan niet 2x gekozen worden
  - volgorde waarin personen worden aangeduid heeft geen belang
- aantal combinaties berekenen van 5 uit 12
- $C_{12}^{5} = \binom{12}{5} = \frac{12!}{5! \cdot (12-5)!} = \frac{12!}{5! \cdot 7!}$

$$= \frac{12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7!}{5! \cdot 7!} = \frac{12 \cdot 11 \cdot 10 \cdot 9 \cdot 8}{5!} = 792$$

- quizploeg kan op 792 verschillende manieren worden samengesteld

Toepassing in Statistiek

- Binomium van Newton is gebaseerd op Driehoek van Pascal:

$$1a + 1b$$

$$1a^2 + 2ab + 1b^2$$

$$1a^3 + 3a^2b + 3ab^2 + 1b^3$$

$$1a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + 1b^4$$

Toepassing in Statistiek

- Driehoek van Pascal bestaat uit binomiaalcoëfficiënten:

$$\binom{1}{0} \quad \binom{1}{1}$$

$$\binom{2}{0} \quad \binom{2}{1} \quad \binom{2}{2}$$

$$\binom{3}{0} \quad \binom{3}{1} \quad \binom{3}{2} \quad \binom{3}{3}$$

$$\binom{4}{0} \quad \binom{4}{1} \quad \binom{4}{2} \quad \binom{4}{3} \quad \binom{4}{4}$$

Toepassing in Statistiek

- vb. $(a+b)^4 = \binom{4}{0} a^4 b^0 + \binom{4}{1} a^3 b^1 + \binom{4}{2} a^2 b^2 + \binom{4}{3} a^1 b^3 + \binom{4}{4} a^0 b^4$

$$= \frac{4!}{0!(4-0)!} a^4 b^0 + \frac{4!}{1!(4-1)!} a^3 b^1 + \frac{4!}{2!(4-2)!} a^2 b^2 + \frac{4!}{3!(4-3)!} a^1 b^3 + \frac{4!}{4!(4-4)!} a^0 b^4$$

$$= a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$$

Toepassing in Statistiek

Oefenmoment:

vb. $(a+b)^6 = ?$

Toepassing in Statistiek

Binomium van Newton

- maakt gebruik van binomiaalcoëfficiënten
- **Formule:**

$$(a+b)^n = \binom{n}{0} a^n b^0 + \binom{n}{1} a^{n-1} b^1 + \ldots + \binom{n}{n-1} a^1 b^{n-1} + \binom{n}{n} a^0 b^n$$

$$= \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

- vb. $(a+b)^4 = \binom{4}{0} a^4 b^0 + \binom{4}{1} a^3 b^1 + \binom{4}{2} a^2 b^2 + \binom{4}{3} a^1 b^3 + \binom{4}{4} a^0 b^4$

Frequentietabellen

vb. Spoed UZA: gedurende 14 dagen noteert men per dag hoeveel armbreuken er zijn:

$4, 3, 1, 1, 2, 6, 1, 1, 1, 2, 2, 1, 2, 3$ → tabelvorm?

Armbreuken	Aantal dagen
1	6
2	4
3	2
4	1
5	0
6	1

Armbreuken = uitkomsten = x_i

Aantal dagen = absolute frequenties = f_i

Frequentietabellen

Wat is het gemiddeld aantal armbreuken per dag?

Armbreuken (x_i)	Aantal dagen (f_i)
1	6
2	4
3	2
4	1
5	0
6	1

$$\overline{x} = \frac{6 \cdot 1 + 4 \cdot 2 + 2 \cdot 3 + 1 \cdot 4 + 0 \cdot 5 + 1 \cdot 6}{6 + 4 + 2 + 1 + 0 + 1} = \frac{30}{14} = 2.14$$

→ gemiddeld 2 (à 3) armbreuken per dag

Frequentietabellen

Algemeen: (zie ook slides woe)

x_i	f_i
1	6
2	4
3	2
4	1
5	0
6	1

Steekproefgemiddelde:

$$\overline{x} = \frac{\sum_{i=1}^{k} f_i \cdot x_i}{n} = \frac{\sum_{i=1}^{k} f_i \cdot x_i}{\sum_{i=1}^{k} f_i}$$

Frequentietabellen

Mediaan?

x_i	1	2	3	4	5	6
f_i	6	4	2	1	0	1

Rangschik de gegevens van klein naar groot en neem de middelste waarneming (wn)

- middelste? $\frac{n+1}{2}$-de geordende wn
- $\frac{14+1}{2}$-de geordende wn = 7.5^e geordende wn
- neem gemiddelde van 7^e en 8^e waarneming
- 7^e wn = 2 en 8^e wn = 2
- Me = 2

Frequentietabellen

Hoe sterk varieert het dagelijks aantal armbreuken?

Armbreuken	Aantal dagen
1	6
2	4
3	2
4	1
5	0
6	1

→ spreidingsmaat berekenen
Hoe ver liggen de uitkomsten verspreid rond het gemiddelde van 2 (à 3) armbreuken per dag?
→ standaardafwijking

Frequentietabellen

steekproefvariantie: $s^2 = \frac{1}{n-1} \sum_{i=1}^{k} f_i \cdot (x_i - \overline{x})^2$
→ $\sqrt{s^2}$ is gelijk aan steekproefstandaardafwijking

Opmerking:

Beter gebruik maken van alternatieve formule!!
(minder rekenwerk, zie ook slides woe)

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{k}\left(f_i \cdot x_i^2\right) - n \cdot \overline{x}^{\,2}\right] \qquad \rightarrow \sqrt{s^2}$$

k = aantal *verschillende* waarnemingen
n = totaal aantal waarnemingen = $\sum_{i=1}^{k} f_i$
$\overline{x}$ = gemiddelde

Hoe sterk varieert het dagelijks aantal armbreuken?

x_i	f_i	x_i^2	$f_i \cdot x_i^2$
1	6	1	6
2	4	4	16
3	2	9	18
4	1	16	16
5	0	25	0
6	1	36	36

$k = 6$; $n = 14$; $\overline{x} = 2.14$

$s^2 = \frac{1}{14-1}\left[\sum_{i=1}^{6}\left(f_i \cdot x_i^2\right) - 14 \cdot (2.14)^{\,2}\right]$

$s^2 = \frac{1}{13}\left[\left(6+16+18+16+0+36\right) - 14 \cdot (2.14)^{\,2}\right]$

$s^2 = 2.145 \rightarrow s = \sqrt{2.145} = 1.46$

Hoeveel dagen waren er maximum 3 armbreuken?

x_i	f_i
1	6
2	4
3	2
4	1
5	0
6	1

$\rightarrow$ nagaan hoe vaak er 1 of 2 of 3 armbreuken waren
$\rightarrow$ absolute frequenties optellen: $6 + 4 + 2 = 12$
$\rightarrow$ bij 12 van de geobserveerde dagen waren er maximum 3 armbreuken

Algemeen:

x_i	f_i	F_i
1	6	6
2	4	10
3	2	12
4	1	13
5	0	13
6	1	14

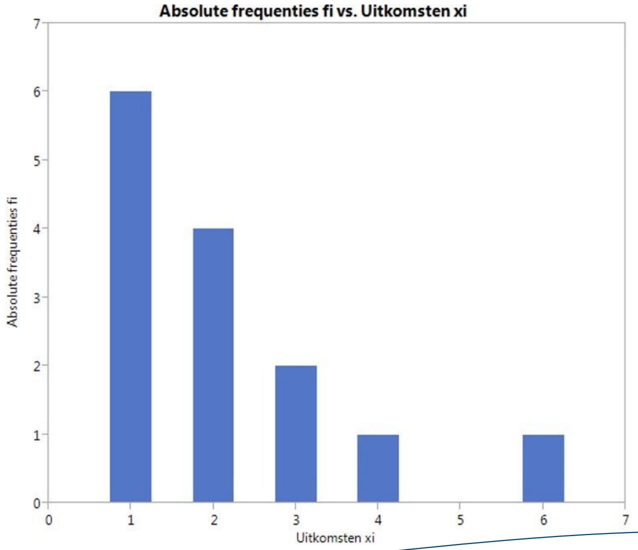
Cumulatieve absolute frequentie = F_i
vb. $F_3 = f_1 + f_2 + f_3 = 6 + 4 + 2 = 12$
vb. $F_4 = f_1 + f_2 + f_3 + f_4 = F_3 + f_4 = 12 + 1 = 13$
enz.

Grafische voorstelling?

x_i	f_i
1	6
2	4
3	2
4	1
5	0
6	1

$\rightarrow$ staafdiagram

Staafdiagram:



vb. besparingsmaatregelen: HR-verantwoordelijke noteert het bruto weekloon van 65 werknemers

Loon (euro)	Aantal wn
[500; 600[	8
[600; 700[	10
[700; 800[	16
[800; 900[	14
[900; 1000[	10
[1000; 1100[	7

→ Gemiddeld bruto weekloon?
→ Hoe rekenen met disjuncte klassen?
→ Gebruik de klassecentra!

Gemiddeld bruto weekloon:

Loon (euro)	Klassecentrum x_i^c	Aantal wn f_i
[500; 600[	550	8
[600; 700[	650	10
[700; 800[	750	16
[800; 900[	850	14
[900; 1000[	950	10
[1000; 1100[	1050	7

$$\overline{x} = \frac{8\cdot550+10\cdot650+16\cdot750+14\cdot850+10\cdot950+7\cdot1050}{8+10+16+14+10+7} = \frac{51650}{65} = 794.62$$

→ gemiddeld bruto weekloon is 794.62 euro

Algemeen:

Klasse	x_i^c	f_i
[500; 600[	550	8
[600; 700[	650	10
[700; 800[	750	16
[800; 900[	850	14
[900; 1000[	950	10
[1000; 1100[	1050	7

Steekproefgemiddelde (o.b.v. klassecentra):

$$\overline{x} = \frac{\sum_{i=1}^{k} f_i \cdot x_i^c}{n} = \frac{\sum_{i=1}^{k} f_i \cdot x_i^c}{\sum_{i=1}^{k} f_i}$$

Frequentietabellen

Steekproefstandaardafwijking?

Klasse	x_i^c	f_i
[500; 600[	550	8
[600; 700[	650	10
[700; 800[	750	16
[800; 900[	850	14
[900; 1000[	950	10
[1000; 1100[	1050	7

$\rightarrow s = \sqrt{\frac{1}{n-1}\left[\sum_{i=1}^{k}\left(f_i \cdot (x_i^c)^2\right) - n \cdot \overline{x}^2\right]}$

met $\overline{x}$ = gemiddelde (o.b.v. klassecentra)

Frequentietabellen

Uitwerking steekproefstandaardafwijking:

Klasse	x_i^c	f_i	$(x_i^c)^2$
[500; 600[	550	8	302500
[600; 700[	650	10	422500
[700; 800[	750	16	562500
[800; 900[	850	14	722500
[900; 1000[	950	10	902500
[1000; 1100[	1050	7	1102500

$k = 6$; $n = 65$; $\overline{x} = 794.62$

$\rightarrow s = \sqrt{\frac{1}{65-1}\left[\sum_{i=1}^{6}\left(f_i \cdot (x_i^c)^2\right) - 65 \cdot (794.62)^2\right]}$

$\rightarrow s = 151.05$

Frequentietabellen

Hoeveel procent van wn verdient tussen 600 euro (incl) en 700 euro (excl) bruto per week?

Klasse	x_i^c	f_i
[500; 600[	550	8
[600; 700[	650	10
[700; 800[	750	16
[800; 900[	850	14
[900; 1000[	950	10
[1000; 1100[	1050	7

$\rightarrow$ relatieve frequentie $= f_i^* = \frac{10}{65} = 0.1538$

$\rightarrow$ iets meer dan 15% van wn verdient tussen 600 en 700 euro bruto per week

Frequentietabellen

Opdacht: Bereken waaraan ? gelijk is en interpreteer.

Klasse	x_i^c	f_i	f_i^*
[500; 600[	550	8	...
[600; 700[	650	10	...
[700; 800[	750	16	?
[800; 900[	850	14	...
[900; 1000[	950	10	...
[1000; 1100[	1050	7	...

$\rightarrow ? = \frac{16}{8+10+16+14+10+7} = \frac{16}{65} = 0.2462$

$\rightarrow \pm$ 25% van de werknemers verdient tussen 700 en 800 euro bruto per week

Alle relatieve frequenties f_i^*:

Klasse	x_i^c	f_i	f_i^*
[500;600[	550	8	0.1231
[600;700[	650	10	0.1538
[700;800[	750	16	0.2462
[800;900[	850	14	0.2154
[900;1000[	950	10	0.1538
[1000;1100[	1050	7	0.1077

→ Som van alle relatieve frequenties?

→ $\sum_{i=1}^{k} f_i^* = 100\% = 1$

Hoeveel procent van wn verdient minder dan 900 euro bruto per week?

Klasse	x_i^c	f_i	f_i^*
[500;600[	550	8	0.1231
[600;700[	650	10	0.1538
[700;800[	750	16	0.2462
[800;900[	850	14	0.2154
[900;1000[	950	10	0.1538
[1000;1100[	1050	7	0.1077

→ $0.1231 + 0.1538 + 0.2462 + 0.2154 = 0.7385$

→ OF: $\frac{8+10+16+14}{65} = \frac{48}{65} = 0.7385$

→ bijna 74% van wn verdient minder dan 900 euro

0.7385 is een cumulatieve relatieve frequentie F_i^*

Klasse	x_i^c	f_i	f_i^*	F_i^*
[500;600[	550	**8**	0.1231	...
[600;700[	650	**10**	0.1538	...
[700;800[	750	**16**	0.2462	...
[800;900[	850	**14**	0.2154	0.7385
[900;1000[	950	10	0.1538	...
[1000;1100[	1050	7	0.1077	...

$0.1231 + 0.1538 + 0.2462 + 0.2154 = 0.7385$

$\frac{\mathbf{8+10+16+14}}{65} = \frac{48}{65} = 0.7385$

Oefening: Wat betekent het rode getal? Hoe werd deze waarde berekend?

Klasse	x_i^c	f_i	f_i^*	F_i^*
[500;600[	550	8	0.1231	...
[600;700[	650	10	0.1538	...
[700;800[	750	16	0.2462	0.5231
[800;900[	850	14	0.2154	...
[900;1000[	950	10	0.1538	...
[1000;1100[	1050	7	0.1077	...

Frequentietabellen

Alle cumulatieve relatieve frequenties F_i^*:

Klasse	x_i^c	f_i	f_i^*	F_i^*
[500; 600[	550	8	0.1231	0.1231
[600; 700[	650	10	0.1538	0.2769
[700; 800[	750	16	0.2462	0.5231
[800; 900[	850	14	0.2154	0.7385
[900; 1000[	950	10	0.1538	0.8923
[1000; 1100[	1050	7	0.1077	1

→ is laatste cumulatieve relatieve frequentie altijd gelijk aan 1?

Frequentietabellen

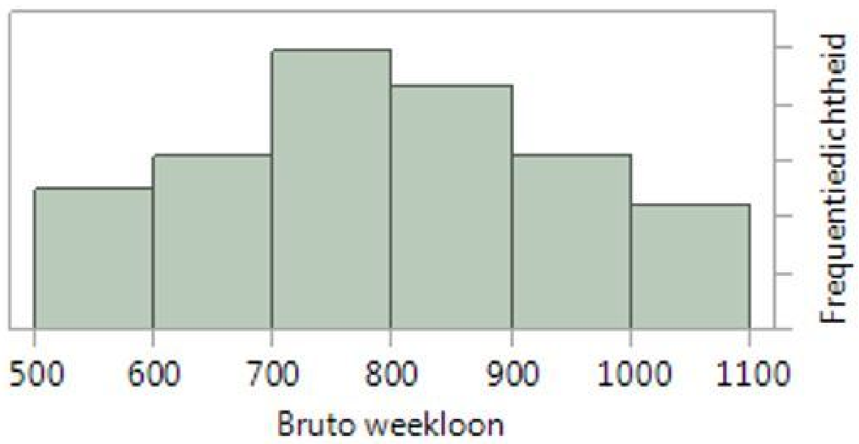
Grafische voorstelling?

Klasse	x_i^c	f_i
[500; 600[	550	8
[600; 700[	650	10
[700; 800[	750	16
[800; 900[	850	14
[900; 1000[	950	10
[1000; 1100[	1050	7

→ histogram
→ frequentiepolygoon

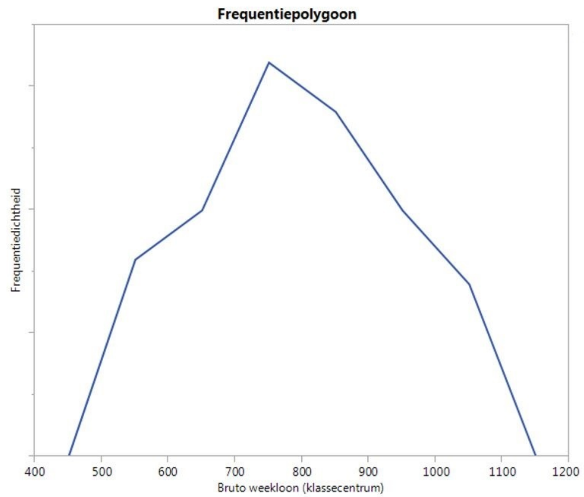
Frequentietabellen

Histogram:



Frequentietabellen

Frequentiepolygoon:



Frequentietabellen

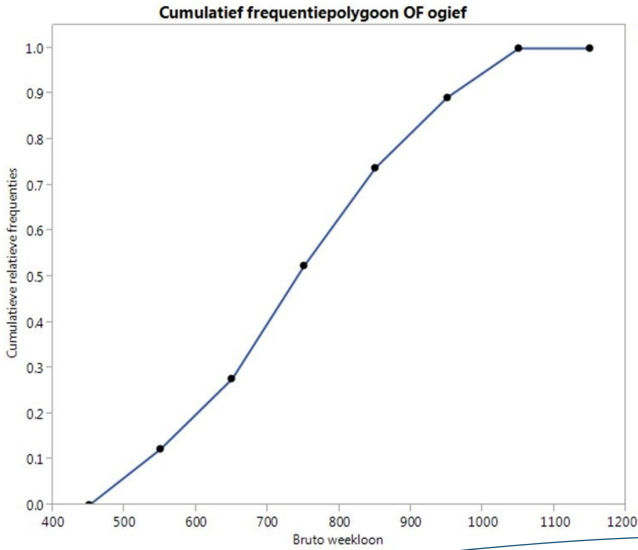
Cumulatieve relatieve frequenties voorstellen?

Klasse	x_i^c	f_i	f_i^*	F_i^*
[500; 600[	550	8	0.1231	0.1231
[600; 700[	650	10	0.1538	0.2769
[700; 800[	750	16	0.2462	0.5231
[800; 900[	850	14	0.2154	0.7385
[900; 1000[	950	10	0.1538	0.8923
[1000; 1100[	1050	7	0.1077	1

→ cumulatief frequentiepolygoon of ogief

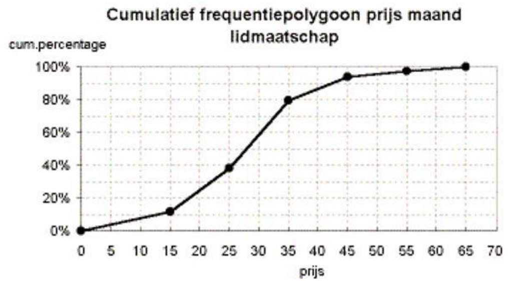
Frequentietabellen

Cumulatief frequentiepolygoon of ogief:



Frequentietabellen

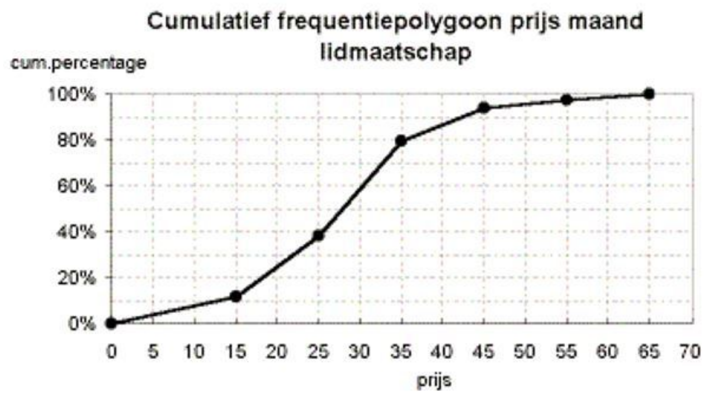
vb. ogief van maandprijs lidmaatschap:



→ (25, 40%) is een punt op het ogief. Wat betekent dat in woorden? 40% van de mensen betaalt 25 euro of minder per maand voor het lidmaatschap

Frequentietabellen

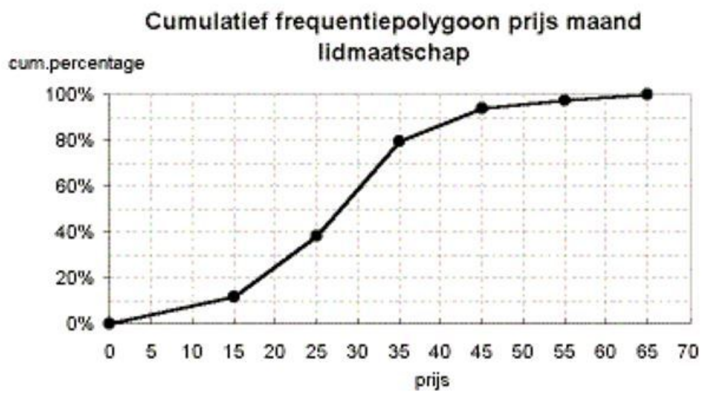
vb. ogief van maandprijs lidmaatschap:



→ Hoeveel procent van mensen betaalt maximum 35 euro? 80%

Frequentietabellen

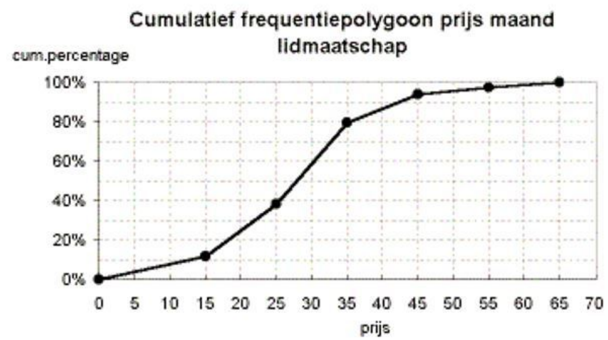
vb. ogief van maandprijs lidmaatschap:



→ Hoeveel procent van mensen betaalt meer dan 25 euro? 60%

Frequentietabellen

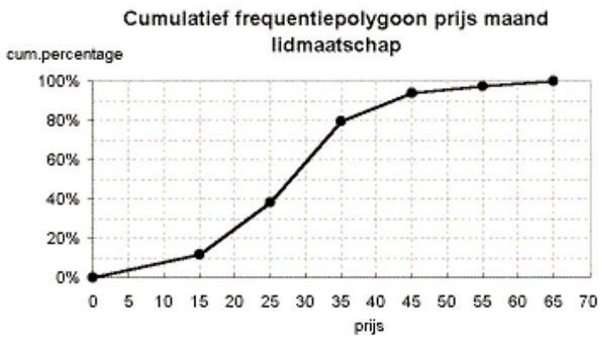
Hoeveel bedraagt de mediaan?



→ Ga na welke prijs er overeenkomt met 50%
→ De mediaan bedraagt ± 28 euro
→ De helft betaalt max 28 euro, de andere helft meer dan 28 euro

Frequentietabellen

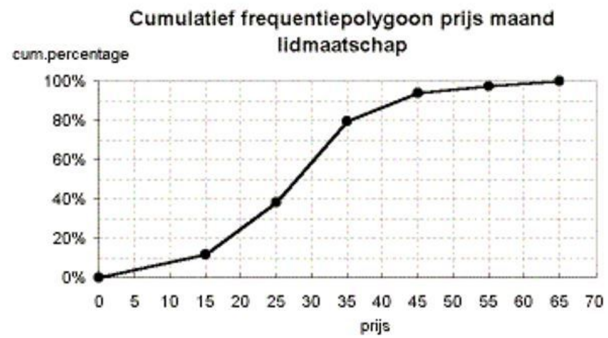
Hoeveel bedraagt het 1ste kwartiel?



→ Ga na welke prijs er overeenkomt met 25%
→ Het 1ste kwartiel bedraagt 20 euro
→ 25% van mensen betaalt max 20 euro, de overige 75% betaalt meer dan 20 euro

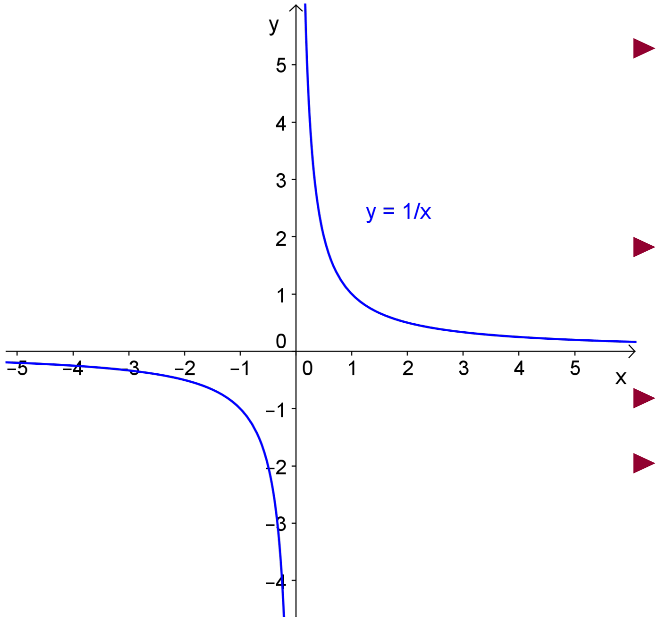
Frequentietabellen

Hoeveel bedraagt het 3de kwartiel?



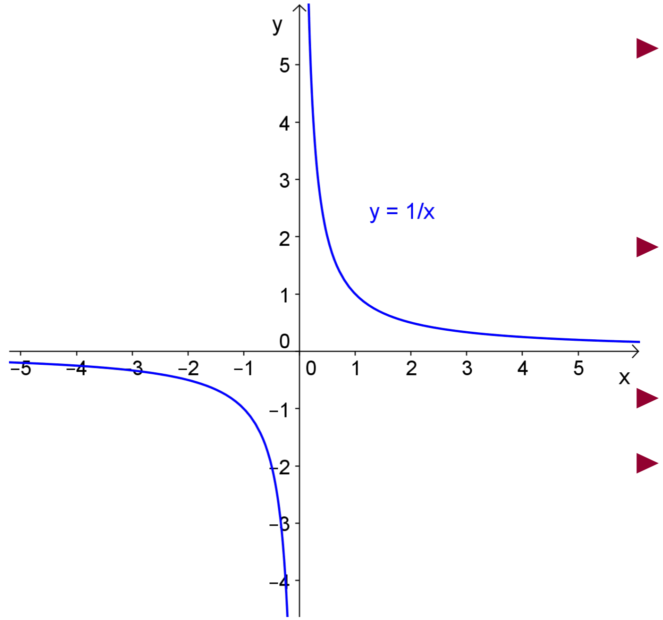
→ Ga na welke prijs er overeenkomt met 75%
→ Het 3de kwartiel bedraagt ± 33 euro
→ 75% van mensen betaalt max 33 euro, de overige 25% betaalt meer dan 33 euro

Limiet



- als we op de x-as naar rechts gaan, dan nadert de hoogte naar 0
- als x naar oneindig gaat, gaat $f(x)$ naar 0:
- $\lim_{x \to +\infty} f(x) = 0$
- 'de limiet van $f(x)$ voor x gaande naar $+\infty$ is 0'

Limiet



- als we op de x-as naar links gaan, dan nadert de hoogte naar 0
- als x naar min oneindig gaat, gaat $f(x)$ naar 0:
- $\lim_{x \to -\infty} f(x) = 0$
- 'de limiet van $f(x)$ voor x gaande naar $-\infty$ is 0'

Toepassing in Statistiek

Empirische kansdefinitie

- vb. dobbelsteen opgooien: hoe groot is de kans om 5 te gooien?
- experiment:

aantal worpen (n)	aantal keer 5 (f_i)	f_i/n
10	1	0.1
100	15	0.15
1000	162	0.162
10 000	1 665	0.1665
100 000	16 666	0.16666
...	...	...
$+\infty$		1/6

Toepassing in Statistiek

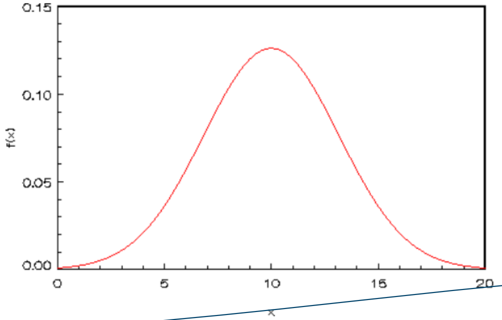
- Als we het aantal pogingen n laten toenemen naar $+\infty$ zal de proportie $\frac{f_i}{n}$ naderen naar $\frac{1}{6}$
- Empirische kansdefinitie

$$P(A) = \lim_{n \to +\infty} \frac{f_i}{n}$$

- In het voorbeeld:
  $P(A) = P(5 \text{ gooien}) =$ de kans om 5 te gooien met deze dobbelsteen $= \frac{1}{6}$

Kansdichtheid

- vb. X = gewicht van peuters
- X is een ‘continue’ kansvariabele
- Continue kansvariabelen hebben een kansdichtheid = positieve functie waarbij de oppervlakte tussen de x-as en de grafiek van de kansdichtheid 1 is
- vb.



Kans - Bepaalde integraal

- Kansdichtheid van X: om kansen waarin X voorkomt te bepalen
- Notatie: $f_X(x)$ is de kansdichtheid van X
- $P(a \leq X \leq b)$

  = **kans** dat X een waarde aanneemt tussen a en b

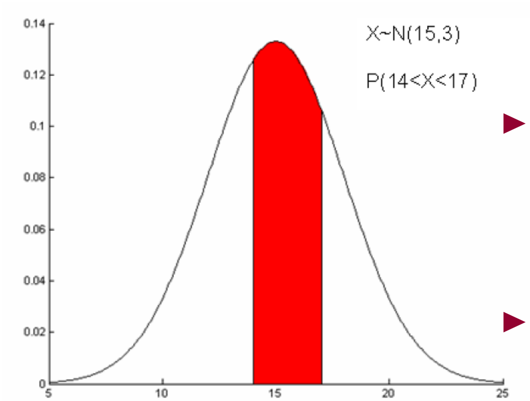
  = **oppervlakte** onder de kansdichtheid van X, tussen a en b

  = $\int_a^b f_X(x)\,\mathrm{d}x$ (bepaalde integraal van f_X voor x gaande van a tot b)

$\rightarrow$ berekenen via tabellen of GRM

Kans - Bepaalde integraal

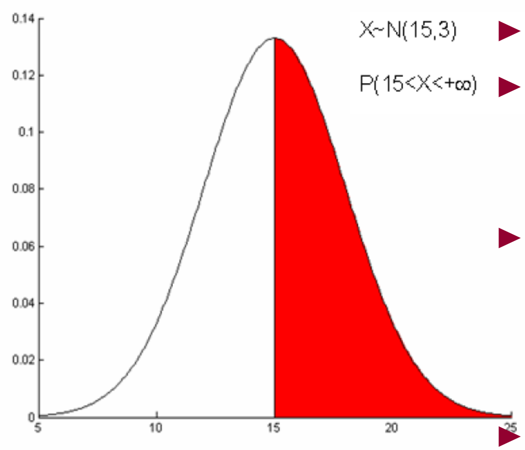
vb.



- $P(14 < X < 17)$
- kans dat X een waarde aanneemt tussen 14 en 17
- oppervlakte onder de kansdichtheid van X, tussen 14 en 17
- $P(14 < X < 17)$ $= 0.3781$ (normalcdf(14, 17, 15, 3))

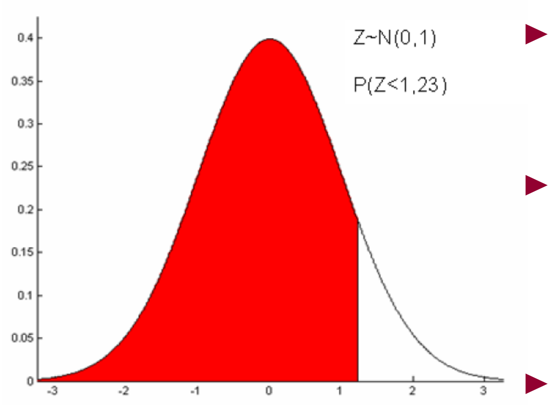
Kans - Bepaalde integraal

vb.



- $P(X > 15)$
- $P(15 < X < +\infty)$
- kans dat X een waarde aanneemt groter dan 15
- oppervlakte onder de kansdichtheid van X, tussen 15 en $+\infty$
- $P(X > 15) = 0.5$

vb.



- $P(Z < 1.23)$
- $P(-\infty < Z < 1.23)$
- kans dat Z een waarde aanneemt kleiner dan 1.23
- oppervlakte onder de kansdichtheid van Z, tussen $-\infty$ en 1.23
- $P(Z < 1.23) = 0.8907$ (normalcdf$(-10^{99}, 1.23)$)