



## Faculteit Militaire Wetenschappen

Gegevens student	
Naam:	
Peoplesoftnummer:	
Klas:	
Handtekening:	

Algemeen			
Vak:	Statistiek deel 2 (tweede kans)	Vakcode:	STA#2
Datum:	13 november 2025	Tijdsduur:	9:00-12:00
Examinator:	Dr. ir. D.A.M.P. Blom	Aantal pagina's:	4
Peer-review:	Dr. ir. B. Westerweel	Aantal opgaven:	4

Algemene instructies
<ul style="list-style-type: none"><li>- Alle antwoorden dienen gemotiveerd te worden. Indien u een deelopgave niet kunt oplossen en het antwoord in vervolgvragen nodig hebt, probeer uit te gaan van een redelijke fictieve waarde.</li><li>- Rond je antwoorden waar nodig af op vier decimalen.</li><li>- U mag een grafische rekenmachine gebruiken zonder CAS (Computer Algebra Systeem).</li><li>- Antwoorden, in welke vorm dan ook, mogen de zaal niet verlaten.</li><li>- Vermeld op elk antwoordvel je naam, Peoplesoft-nummer en maak een nummering van je antwoordvellen.</li><li>- Iedere vorm van mobiele (potentiële) datadragers (telefoon, smartwatch, etc.) of andere vormen om te frauderen (bv. communicatieapparatuur) zijn niet toegestaan gedurende de gehele duur van het tentamen en mogen ook niet in het lokaal meegebracht worden of zijn uitgeschakeld en ingeleverd.</li><li>- Schrijf leesbaar ter voorkoming van misverstanden bij de beoordeling van uw werk. Indien uw antwoord niet leesbaar is, wordt uw antwoord fout gerekend.</li><li>- Toiletbezoek tijdens het tentamen vindt enkel plaats na toestemming van de examinerator.</li><li>- Lever bij het verlaten van de zaal, kladpapier, tentamenopgaven en andere tentamen-gerelateerde documenten in bij de examinerator.</li></ul>

**Cijferberekening / cesuur**

- Het eindcijfer voor het vak Statistiek wordt voor 50% bepaald door dit tentamen.
- Het tentamen is opgebouwd uit 4 open vragen. Bij iedere (sub)vraag is het aantal te behalen punten tussen haakjes aangegeven. In totaal kunt u 100 punten verdienen.
- Het tentamencijfer wordt bepaald door het totaal aantal punten te delen door 10. Het tentamencijfer moet minimaal een 5,0 zijn om de cursus Statistiek met succes af te ronden.

**Procedure na het tentamen**

- De cijfers van dit tentamenonderdeel worden in principe binnen 10 werkdagen na de afname bekend gemaakt.
- Met vragen over de beoordeling kunt u tot 10 werkdagen na bekendmaking van de cijfers terecht bij de cursuscoördinator.

Veel succes!

# Formuleblad Statistiek (2024-2025)

## Statistiek deel 1

**Steekproefgemiddelde (gegeven een steekproef met  $n$  uitkomsten  $x_1, x_2, \dots, x_n$ )**

$$\bar{x} = \frac{\sum_i x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

**Steekproefvariantie en steekproefstandaardafwijking:**

$$s^2 = \frac{\sum_i (x_i - \bar{x})^2}{n - 1} = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n - 1}$$
$$s = \sqrt{s^2} = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n - 1}}$$

**Rekenregels kansrekening:**

$$P(A \text{ of } B) = P(A) + P(B) - P(A \text{ en } B) \quad (\text{optelregel})$$

$$P(B) = 1 - P(\text{niet } B) \quad (\text{complementregel})$$

$$P(A | B) = \frac{P(A \text{ en } B)}{P(B)} \quad (\text{conditionele kansen})$$

**Discrete en continue kansverdelingen:**

	Discrete kansvariabelen	Continue kansvariabelen
<b>Uitkomstenruimte:</b>	Eindig / aftelbaar oneindig	Overaftelbaar oneindig
<b>Toepassingen:</b>	Tellen / categoriseren	Metten
<b>Kansbegrip:</b>	Kansfunctie $p(k) = P(X = k)$	Kansdichtheidsfunctie $f(x)$
<b>CDF:</b>	$F(k) = P(X \leq k) = \sum_{\ell: \ell \leq k} p(\ell)$	$F(x) = P(X \leq x) = \int_{-\infty}^x f(y) dy$
<b>Verwachtingswaarde:</b>	$E[X] = \sum_k k \cdot P(X = k)$	$E[X] = \int x \cdot f(x) dx$
<b>Variantie:</b>	$\text{Var}(X) = \sum_k (k - E[X])^2 \cdot P(X = k)$	$\text{Var}(X) = \int (x - E[X])^2 \cdot f(x) dx$
<b>Standaardafwijking:</b>	$\sigma(X) = \sqrt{\text{Var}(X)}$	$\sigma(X) = \sqrt{\text{Var}(X)}$

**Speciale kansverdelingen:**

- $X \sim \text{Binomiaal}(n, p)$ : tellen van aantal successen bij onafhankelijke kansexperimenten met twee uitkomsten (Bernoulli-experimenten): succes / mislukking.

**Parameters:** het aantal Bernoulli-experimenten  $n$  en de succeskans per experiment  $p$ .

- $X \sim \text{Poisson}(\lambda \cdot t)$ : tellen van aantal “gebeurtenissen” in een “interval” van tijd / ruimte.

**Parameters:** het gemiddelde aantal gebeurtenissen  $\lambda$  per meeteenheid (tijd / ruimte) en het aantal meeteenheden  $t$ .

→ Voorbeeld: bij de meeteenheid van een dag bestaat een week uit  $t = 7$  meeteenheden.

- $T \sim \text{Exponentieel}(\lambda)$ : meten van de tijd / ruimte tot de volgende gebeurtenis.

**Parameter:** het gemiddelde aantal gebeurtenissen  $\lambda$  per meeteenheid (tijd / ruimte).

## Verwachtingswaarde en variantie van veelgebruikte kansverdelingen:

Verdeling	Kans(dichtheids)functie	CDF	$E(X)$	$\text{Var}(X)$
<b>Discreet</b>				
Uniform( $a, b$ )	$p(k) = \frac{1}{b-a+1}$ ( $k = a, a+1, \dots, b$ )	$F(k) = \begin{cases} 0 & x < a \\ \frac{k-a+1}{b-a+1} & a \leq k < b \\ 1 & k \geq b \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a+1)^2-1}{12}$
Binomiaal( $n, p$ )	$p(k) = \binom{n}{k} p^k (1-p)^{n-k}$	$F(k) = \sum_{i=0}^k \binom{n}{i} p^i (1-p)^{n-i}$	$np$	$np(1-p)$
Poisson( $\lambda$ )	$p(k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$	$F(k) = \sum_{i=0}^k e^{-\lambda} \cdot \frac{\lambda^i}{i!}$	$\lambda$	$\lambda$
<b>Continuous</b>				
Uniform( $a, b$ )	$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{elders.} \end{cases}$	$F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x < b \\ 1, & x \geq b \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponentieel( $\lambda$ )	$f(x) = \lambda e^{-\lambda x}, x \geq 0$	$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

## Veelgebruikte functies op de grafische rekenmachine

Type vraag	TI-84 Plus	Casio
<b>Continue kansverdeling (willekeurig)</b>		
$P(a \leq X \leq b)$	$\int_a^b f(x) dx$	$\int_a^b f(x) dx$
$X \sim \text{Binomiaal}(n, p)$		
$P(X = k)$	binompdf( $n, p, k$ )	BinomialPD( $k, n, p$ )
$P(X \leq k)$	binomcdf( $n, p, k$ )	BinomialCD( $k, n, p$ )
$X \sim N(\mu, \sigma)$		
$P(a \leq X \leq b)$	normalcdf( $a, b, \mu, \sigma$ )	NormalCD( $a, b, \sigma, \mu$ )
Grenswaarde $g$ zodat $P(X \leq g) = p$ ?	invNorm( $p, \mu, \sigma$ )	InvNormCD(tail=left, $p, \sigma, \mu$ )
$X \sim \text{Poisson}(\lambda)$		
$P(X = k)$	poissonpdf( $\lambda, k$ )	PoissonPD( $k, \lambda$ )
$P(X \leq k)$	poissoncdf( $\lambda, k$ )	PoissonCD( $k, \lambda$ )

## $z$ -score:

$$z = \frac{x - \mu}{\sigma}$$

**Centrale limietstelling:** Gegeven  $n$  kansvariabelen  $X_1, X_2, \dots, X_n$  die onderling onafhankelijk zijn en dezelfde kansverdeling hebben met een verwachtingswaarde  $\mu$  en standaardafwijking  $\sigma$ , dan geldt (bij benadering) dat

- de som  $\sum X = X_1 + X_2 + \dots + X_n$  normaal verdeeld is met verwachtingswaarde  $n \cdot \mu$  en standaardafwijking  $\sqrt{n} \cdot \sigma$ .
- het gemiddelde  $\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$  normaal verdeeld is met verwachtingswaarde  $\mu$  en standaardafwijking  $\frac{\sigma}{\sqrt{n}}$ .

---

## Statistiek deel 2:

### Betrouwbaarheidsintervallen voor het gemiddelde $\mu$ ( $\sigma$ bekend)

- $100 \cdot (1 - \alpha)\%$ -betrouwbaarheidsinterval (BI) voor  $\mu$ :

$$z_{\alpha/2} = \text{InvNorm}(\text{opp} = 1 - \alpha/2; \mu = 0; \sigma = 1)$$

$$\left[ \bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}; \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right]$$

- Minimale steekproefomvang voor  $100 \cdot (1 - \alpha)\%$ -BI als  $\mu$  maximaal  $\pm a$  mag afwijken:

$$n \geq \left( \frac{z_{\alpha/2} \cdot \sigma}{a} \right)^2$$

### Betrouwbaarheidsintervallen voor het gemiddelde $\mu$ ( $\sigma$ onbekend)

- $100 \cdot (1 - \alpha)\%$ -betrouwbaarheidsinterval (BI) voor  $\mu$ :

$$t = \text{InvT}(\text{opp} = 1 - \alpha/2; \text{df} = n - 1)$$

$$\left[ \bar{x} - t \cdot \frac{s}{\sqrt{n}}; \bar{x} + t \cdot \frac{s}{\sqrt{n}} \right]$$

- Minimale steekproefomvang voor  $100 \cdot (1 - \alpha)\%$ -BI als  $\mu$  maximaal  $\pm a$  mag afwijken:

$$\text{GR tabel (voor verschillende } n): \frac{s}{\sqrt{n}} \cdot \text{InvT}(\text{opp} = 1 - \alpha/2; \text{df} = n - 1) \leq a$$

NB: zodra  $n \geq 30$ , vallen de normale en de  $t$ -verdeling nagenoeg samen. Je mag dan rekenen met de schatting  $s$  in plaats van de daadwerkelijke (onbekende)  $\sigma$ .

- Onderscheidend vermogen (toets met  $H_0 : \mu = \mu_0$  vs.  $H_1 : \mu \neq \mu_0$ , en gegeven  $\mu = \mu_1$ )

$$1 - \beta = P(\bar{X} \text{ neemt waarde aan in het kritieke gebied} \mid \mu = \mu_1)$$

### Betrouwbaarheidsintervallen voor de binomiale succeskans $p$

**Betrouwbaarheidsinterval voor  $p$  (Clopper-Pearson):** Gegeven een binomiale verdeling met  $n$  Bernoulli-experimenten en onbekende  $p$ , en uitkomst  $k$ .

1. Bereken de succeskans  $p_1$  zodat geldt  $P(X \leq k) = \text{binomcdf}(n; p; k) = \alpha/2$
2. Bereken de succeskans  $p_2$  zodat geldt  $P(X \geq k) = 1 - \text{binomcdf}(n; p; k - 1) = \alpha/2$
3. De berekende waarden voor  $p_1$  en  $p_2$  zijn de grenzen van het Clopper-Pearson interval.

## Hypothesetoetsen

### Stappenplan hypothesetoetsen

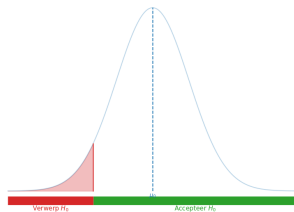
1. Definieer de nulhypothese  $H_0$  en de alternatieve hypothese  $H_1$ .
2. Bepaal het significantieniveau  $\alpha$  (kans op verwerpen van  $H_0$  terwijl  $H_0$  waar is  $\rightarrow$  type-I fout)
3. Verzamel data voor de toetsingsgrootte
4. Bereken de toetsingsgrootte
  - Uitgaande van de nulhypothese  $H_0$  maken we aannames over de kansverdeling van de toetsingsgrootte!
5. Geef een conclusie (met behulp van het kritieke gebied /  $p$ -waarde) en vertaal deze terug naar de originele probleemcontext.

## Drie typen hypothesetoetsen: linkszijdig, tweezijdig, rechtszijdig

### Linkszijdige toets

Kritiek gebied:

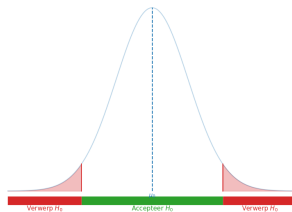
$$(-\infty; g]$$



### Tweezijdige toets

Kritiek gebied:

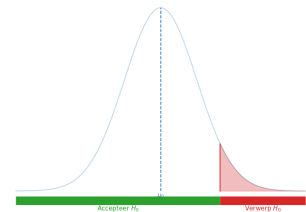
$$(-\infty; g_1] \text{ en } [g_2; \infty)$$



### Rechtszijdige toets

Kritiek gebied:

$$[g; \infty)$$



Kansverdeling (onder $H_0$ )	Linkszijdig	Tweezijdig	Rechtszijdig
$N(\mu; \sigma)$	$g = \text{InvNorm}(\alpha; \mu; \sigma)$	$g_1 = \text{InvNorm}(\text{opp} = \frac{\alpha}{2}; \mu; \sigma)$ $g_2 = \text{InvNorm}(\text{opp} = 1 - \frac{\alpha}{2}; \mu; \sigma)$	$g = \text{InvNorm}(1 - \alpha; \mu; \sigma)$
$t(\text{df})$	$g = \text{InvT}(\alpha; \text{df})$	$g_1 = \text{InvT}(\text{opp} = \frac{\alpha}{2}; \text{df})$ $g_2 = \text{InvT}(\text{opp} = 1 - \frac{\alpha}{2}; \text{df})$	$g = \text{InvT}(1 - \alpha; \text{df})$
Grenzen die met de solver functie moeten worden opgelost:			
$\chi^2(\text{df})$ (chikwadraat)	$\chi^2\text{cdf}(0; g; \text{df}) = \alpha$	$\chi^2\text{cdf}(0; g_1; \text{df}) = \frac{\alpha}{2}$ $\chi^2\text{cdf}(g_2; 10^{99}; \text{df}) = \frac{\alpha}{2}$	$\chi^2\text{cdf}(g; 10^{99}; \text{df}) = \alpha$
$F(\text{df}_A; \text{df}_B)$	$\text{Fcdf}(0; g; \text{df}_A; \text{df}_B) = \alpha$	$\text{Fcdf}(0; g_1; \text{df}_A; \text{df}_B) = \frac{\alpha}{2}$ $\text{Fcdf}(g_2; 10^{99}; \text{df}_A; \text{df}_B) = \frac{\alpha}{2}$	$\text{Fcdf}(g; 10^{99}; \text{df}_A; \text{df}_B) = \alpha$

## $p$ -waardes uitrekenen (gegeven een theoretische en geobserveerde toetsingsgrootheid $T$ en $t$ )

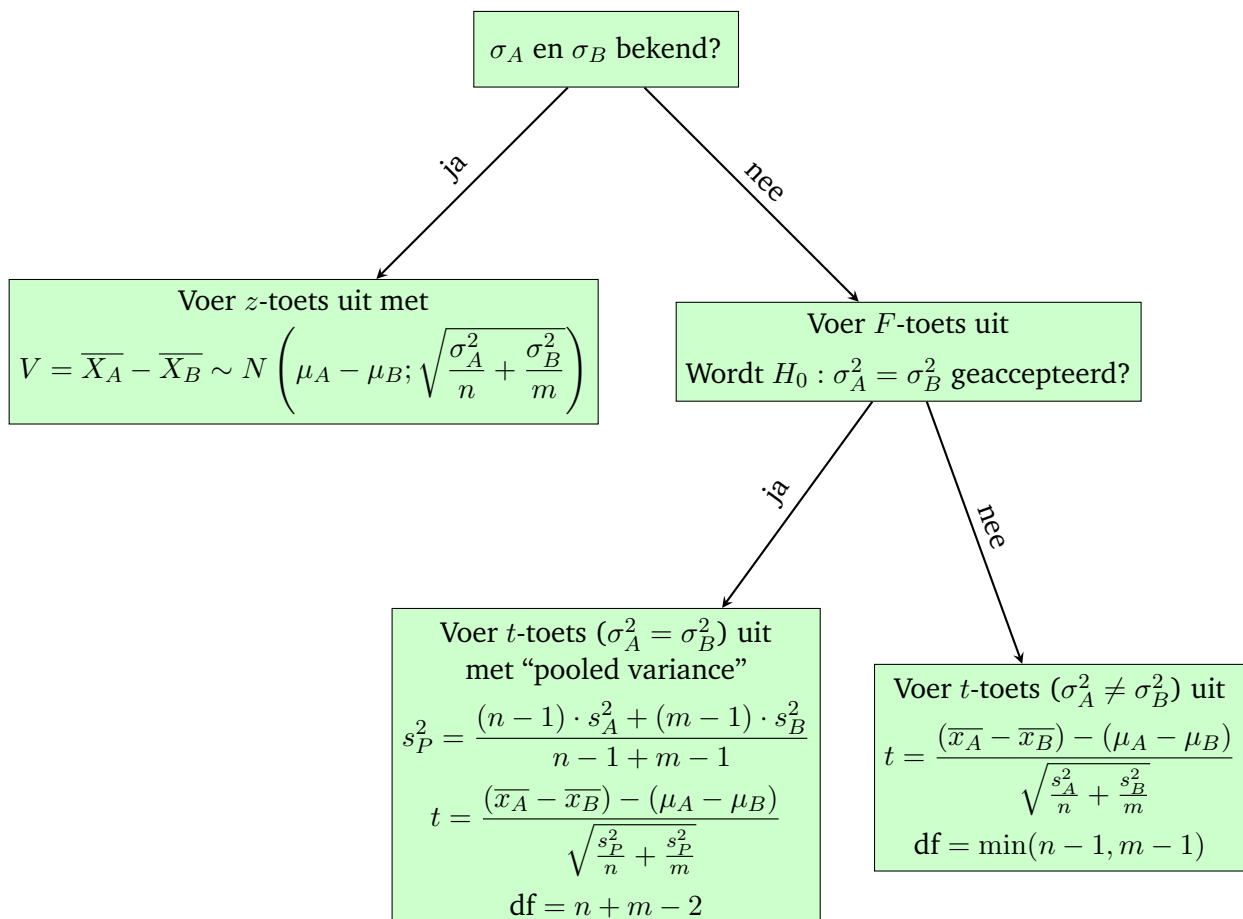
Kansverdeling (onder $H_0$ )	Linkszijdig ( $P(T \leq t)$ )	Rechtszijdig ( $P(T \geq t)$ )
$N(\mu; \sigma)$	$p = \text{normalcdf}(-10^{99}; t; \mu; \sigma)$	$p = \text{normalcdf}(t; 10^{99}; \mu; \sigma)$
$t(\text{df})$	$p = \text{tcdf}(-10^{99}; t; \text{df})$	$p = \text{tcdf}(t; 10^{99}; \text{df})$
$\chi^2(\text{df})$	$p = \chi^2\text{cdf}(0; t; \text{df})$	$p = \chi^2\text{cdf}(t; 10^{99}; \text{df})$
$F(\text{df}_A; \text{df}_B)$	$p = \text{Fcdf}(0; t; \text{df}_A; \text{df}_B)$	$p = \text{Fcdf}(t; 10^{99}; \text{df}_A; \text{df}_B)$

NB: Om met de  $p$ -waarde een conclusie te trekken uit een hypothesetoets vergelijken we de  $p$ -waarde met het significantieniveau  $\alpha$ . Let op: bij tweezijdige toetsen neem je het minimum van de linkszijdige en rechtszijdige  $p$ -waarde en vergelijk je deze met  $\alpha/2$ !

## Soorten toetsen

Soort toets	Toetsingsgrootheid	Kansverdeling (onder $H_0$ )
<b>Toetsen voor het gemiddelde <math>\mu \leq \mu_0</math> of <math>\mu = \mu_0</math> of <math>\mu \geq \mu_0</math></b>		
$z$ -toets ( $\sigma$ bekend)	$\bar{X}$	$N(\mu_0; \frac{\sigma}{\sqrt{n}})$
$t$ -toets ( $\sigma$ onbekend)	$T = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$	$t(df = n - 1)$
<b>Chikwadraattoetsen (<math>\chi^2</math>)</b>		
Onafhankelijkheid	$X^2 = \sum_{i,j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$	$\chi^2(df = (\#rijen-1) \cdot (\#kolommen-1))$
Aanpassing (goodness-of-fit)	$X^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$	$\chi^2(df = (\#categorien-1))$
<b>Verschiltoetsen (op basis van twee populaties <math>A</math> en <math>B</math>)</b>		
$F$ -toets: $\sigma_A^2 = \sigma_B^2$	$F = \frac{S_A^2}{S_B^2}$	$F(df_A, df_B)$
$z$ -toets	$V = \bar{X}_A - \bar{X}_B$	$N\left(\mu_A - \mu_B; \sqrt{\frac{\sigma_A^2}{n} + \frac{\sigma_B^2}{m}}\right)$
$t$ -toets ( $\sigma_A^2 = \sigma_B^2$ )	$T = \frac{(\bar{X}_A - \bar{X}_B) - (\mu_A - \mu_B)}{\sqrt{\frac{s_P^2}{n} + \frac{s_P^2}{m}}}$	$t(df = n + m - 2)$
$t$ -toets ( $\sigma_A^2 \neq \sigma_B^2$ )	$T = \frac{(\bar{X}_A - \bar{X}_B) - (\mu_A - \mu_B)}{\sqrt{\frac{s_A^2}{n} + \frac{s_B^2}{m}}}$	$t(df = \min(n - 1; m - 1))$

## Beslisboom verschiltoetsen



---

## Correlatie en regressie

**Correlatiecoëfficiënt van Pearson:**

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sqrt{(\overline{x^2} - \bar{x}^2) \cdot (\overline{y^2} - \bar{y}^2)}}$$

**Correlatiecoëfficiënt van Spearman:**

$$r_s = 1 - \frac{6 \cdot \sum_i d_i^2}{n^3 - n}$$

**Coëfficiënten van de lineaire regressielijn  $Y = a + b \cdot X$ :**

$$b = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - (\bar{x})^2}$$
$$a = \bar{y} - b \cdot \bar{x}$$

**Schatting van de variantie van de storingsterm  $\varepsilon$ :**

$$s_\varepsilon^2 = \frac{\sum e_i^2}{n - 2} = \frac{\sum (y_i - (a + b \cdot x_i))^2}{n - 2} = \frac{n}{n - 2} \cdot (\overline{y^2} - a \cdot \bar{y} - b \cdot \overline{xy})$$

**$100 \cdot (1 - \alpha)\%$ -betrouwbaarheidsinterval voor de gemiddelde  $Y$  bij een gegeven  $X = x_0$ :**

$$t = \text{InvT}(\text{opp} = 1 - \alpha/2; \text{df} = n - 2)$$

$$s_\mu = s_\varepsilon \cdot \sqrt{\frac{1}{n} \cdot \left(1 + \frac{(x_0 - \bar{x})^2}{\overline{x^2} - \bar{x}^2}\right)}$$

$$[a + b \cdot x_0 - t \cdot s_\mu; a + b \cdot x_0 + t \cdot s_\mu]$$

**$100 \cdot (1 - \alpha)\%$ -betrouwbaarheidsinterval voor  $Y$  bij een gegeven  $X = x_0$ :**

$$t = \text{InvT}(\text{opp} = 1 - \alpha/2; \text{df} = n - 2)$$

$$s_f = s_\varepsilon \cdot \sqrt{1 + \frac{1}{n} \cdot \left(1 + \frac{(x_0 - \bar{x})^2}{\overline{x^2} - \bar{x}^2}\right)}$$

$$[a + b \cdot x_0 - t \cdot s_f; a + b \cdot x_0 + t \cdot s_f]$$



**Opgave 1 (27 punten)** Een onderzoekseenheid van Damen Naval voert een analyse uit op de relatie tussen de snelheid van een marinefregat (in knopen) en haar akoestische signatuur (in decibel). Hiervoor worden 11 fregatten aselekt gekozen en worden de volgende metingen verricht:

Snelheid (knopen)	8.0	10.0	12.0	14.0	16.0	18.0	20.0	22.0	24.0	26.0	28.0
Akoestische signatuur (dB)	116.7	124.0	123.3	123.5	120.6	122.1	122.6	132.2	131.0	133.6	128.2

- 1a [2pt]** Als we een regressie-analyse willen uitvoeren, welke variabele zou dan de afhankelijke variabele  $Y$  zijn en welke variabele zou de onafhankelijke variabele  $X$  zijn?
- 1b [4pt]** Teken het bijbehorende spreidingsdiagram op basis van je antwoord op vraag a. Geef hierbij duidelijk aan welke as welke variabele voorstelt.
- 1c [7pt]** Bereken Pearson's correlatiecoëfficiënt  $r(x, y)$  op basis van de data in de bovenstaande tabel. Wat zegt  $r(x, y)$  over de relatie tussen snelheid en akoestische signatuur?
- 1d [6pt]** Bereken de regressielijn  $Y = a + b \cdot X$  door berekening van de coëfficiënten  $a$  en  $b$ . Bepaal aan de hand van de regressielijn een statistisch verantwoorde voorspelling voor de akoestische signatuur van een fregat dat 11.2 knopen vaart.
- 1e [8pt]** Bereken een 95 %-betrouwbaarheidsinterval voor de gemiddelde akoestische signatuur van fregatten die met een kruissnelheid van 11.2 knopen varen. Rond af op gehele getallen zodanig dat het betrouwbaarheidsniveau gewaarborgd blijft.

**Opgave 2 (27 punten)** Een luchtmachteenheid is een nieuw type radar aan het testen voor het detecteren van vijandelijke drones. Fabrikant Thales claimt dat de radar een succeskans van 70 % heeft om een drone te detecteren (onafhankelijk van andere drones). Om deze claim te testen, worden er 1000 onafhankelijke tests uitgevoerd. In elke test worden vier drones op het systeem afgestuurd en geteld hoeveel van de vier drones gedetecteerd worden. De gegevens zijn weergegeven in onderstaande frequentietabel:

Aantal drones	Frequentie
0	15
1	105
2	290
3	360
4	230

Om de claim van de fabrikant te toetsen, wordt een chikwadraat aanpassingstoets uitgevoerd.

**2a [4pt]** Welke kansverdeling volgt het aantal gedetecteerde drones in één enkele test met een radar? Geef daarnaast specifieke waarden van de bijbehorende parameters.

**2b [3pt]** Formuleer de nulhypothese  $H_0$  en de alternatieve  $H_1$  van deze hypothesetoets. Wat zou in deze context de betekenis zijn van het verwerpen van de nulhypothese?

**2c [3pt]** Bereken de verwachte (“expected”) frequenties van het aantal gedetecteerde drones op basis van de eerder genoemde 1000 onafhankelijke tests, uitgaande van  $H_0$ .

**2d [5pt]** Bereken de toetsingsgrootheid en de  $p$ -waarde op basis van de gegeven frequenties.

**2e [5pt]** Formuleer een conclusie voor deze chikwadraattoets (op basis van een significantieniveau  $\alpha = 0.05$ ) en vertaal deze terug naar de originele probleemcontext. Verklaar deze conclusie aan de hand van de geobserveerde en verwachte frequenties.

**2f [7pt]** Bereken een 95 %-betrouwbaarheidsinterval voor de succeskans  $p$  met de Clopper-Pearson methode. Wat zegt dit over de claim van Thales van 70 % kans op detectie?

**Hint:** gebruik dat 1000 onafhankelijke realisaties van een binomiale kansvariabele met  $n = 4$  in feite neerkomt op een enkele realisatie van één binomiale kansvariabele met  $n = 4000$ .

**Opgave 3 (28 punten)** De Koninklijke Marine is geïnteresseerd in het bepalen van een betrouwbare onderhoudsstrategie voor de maritieme NH90-helikopters. Hiervoor is het belangrijk om data te verzamelen van de *mean time between failures* (MTBF), oftewel de gemiddelde tijd tussen twee faalmomenten van een helikopteronderdeel. Om te onderzoeken wat kritieke onderdelen zijn, worden de *time between failures* (TBF) van de motor ( $X$ ) en rotorbladen ( $Y$ ) van tien NH90-helikopters gemeten.

De volgende data zijn verzameld over de time between failures van motoren en rotorbladen van NH90-helikopters.

TBF motoren (uren)	1185	1175	1195	1180	1195	1190	1185	1215	1175	1205
TBF rotorbladen (uren)	1180	1205	1190	1210	1175	1200	1225	1195	1185	1215

De centrale vraag is nu om te toetsen of de MTBF  $\mu_X$  van de motor significant lager is dan de MTBF  $\mu_Y$  van de rotorbladen, oftewel dat de motor gemiddeld genomen sneller faalt dan de rotorbladen. Voor beide steekproeven kan worden aangenomen dat de tijden tussen faalmomenten normaal verdeeld zijn.

**3a [8pt]** Bepaal voor beide populaties de steekproefgemiddelden ( $x$  en  $y$ ) en de steekproefvarianties ( $s_X^2$  en  $s_Y^2$ ).

**3b [10pt]** Voer een  $F$ -toets uit om te bepalen of de varianties van de times between failures (TBFs) van de twee populaties als gelijk kunnen worden beschouwd ( $\sigma_X^2 = \sigma_Y^2$ ). Gebruik hiervoor een significantieniveau van  $\alpha = 0.05$  en bepaal de toetsuitslag op basis van het kritieke gebied.

**3c [10pt]** Bepaal met behulp van een onafhankelijke  $t$ -toets of de MTBF van de motor significant lager is dan die van de rotorbladen. Gebruik hiervoor opnieuw een significantieniveau van  $\alpha = 0.05$ , en bepaal de toetsuitslag op basis van de  $p$ -waarde.

---

**Opgave 4 (18 punten)** Tijdens patrouilles in oefengebieden kan de Landmacht geconfronteerd worden met verborgen explosieven of boobytraps. Bij een vermoeden van een boobytrap wordt de EOD (Explosieve Opruimingsdienst Defensie) ingeschakeld om de situatie te beoordelen en te neutraliseren indien nodig. De patrouilles testen een nieuwe sensortechnologie waarmee mogelijk sneller een mogelijke boobytrap kan worden gesignaleerd en correct doorgegeven aan de EOD.

In een veldtest wordt het sensorsysteem getest en zijn 18 opeenvolgende detecties gemeten waarbij de reactietijd (in seconden) van het eerste signaal tot de melding aan EOD werd geregistreerd. Historisch ligt de gemiddelde reactietijd bij het oude meldingsprotocol op 11.8 seconden. Met de nieuwe sensoren werd een gemiddelde van 11.2 seconden gemeten, met een standaarddeviatie van 1.6 seconden. De vraag is of deze versnelling statistisch significant is, zodat de Landmacht kan besluiten tot bredere implementatie van de technologie. We mogen aannemen dat de reactietijden normaal verdeeld zijn.

- 4a [3pt]** Formuleer de nulhypothese  $H_0$  en de alternatieve hypothese  $H_1$  om te testen of de nieuwe sensortechnologie een snellere reactietijd biedt dan het oude meldingsprotocol.
- 4b [6pt]** Bereken – onder de nulhypothese  $H_0$  – een 95 %-voorspellingsinterval voor de gemiddelde reactietijd van 18 willekeurige metingen met de nieuwe sensortechnologie.
- 4c [4pt]** Concludeer aan de hand van het voorspellingsinterval van vraag (b) of de gemeten gemiddelde reactietijd van 11.2 seconden met de nieuwe sensortechnologie wijst op een statistisch significante verbetering in reactietijd vergeleken met het oude meldingsprotocol.
- 4d [5pt]** Nog steeds uitgaande van een steekproefgemiddelde van 11.2 seconden en een steekproefstandaardafwijking van 1.6 seconden, hoe groot had de steekproefomvang minimaal moeten zijn om tot een andere conclusie te zijn gekomen?