

Tentamen (Wiskundige) Statistiek Module 5 Statistiek en Analyse

Bachelor 2 Technische Wiskunde

Modulecode: 201400218
Datum: dit is **voorbeeldtentamen B**
Tijd: tijdsduur is 3 uur
Blokcoördinator: Dr. P.K. Mandal
Docent: Dr. K. Poortema

Toegestane middelen: alleen gewone rekenmachine (geen GR)

Bijlagen: standaardnormale tabel,
t-tabel,
chi-kwadraat-tabel,
Tabel F verdeling (1 pagina, $\alpha = 0.05$)
formuleblad Wiskundige Statistiek (2 pagina's)

Normering:

1	2a	2b	2c	2d	3	4	5	6a	6b	7a	7b	8a	8b	8c
2	2	2	2	2	4	4	2	4	2	3	1	1	1	4

Totaal: 36 punten

Opmerking 1:

Voor dit voorbeeldtentamen hebben we zo veel mogelijk gebruik gemaakt van werkcollegeopgaven. Engelstalige opgaven hebben we gemakshalve niet vertaald. De opgaven van de tentamens van studiejaar 2016/2017 zijn in het Nederlands gesteld, met als enige mogelijke uitzondering beschrijvingen van data sets (die onvertaald overgenomen worden).

Opmerking 2:

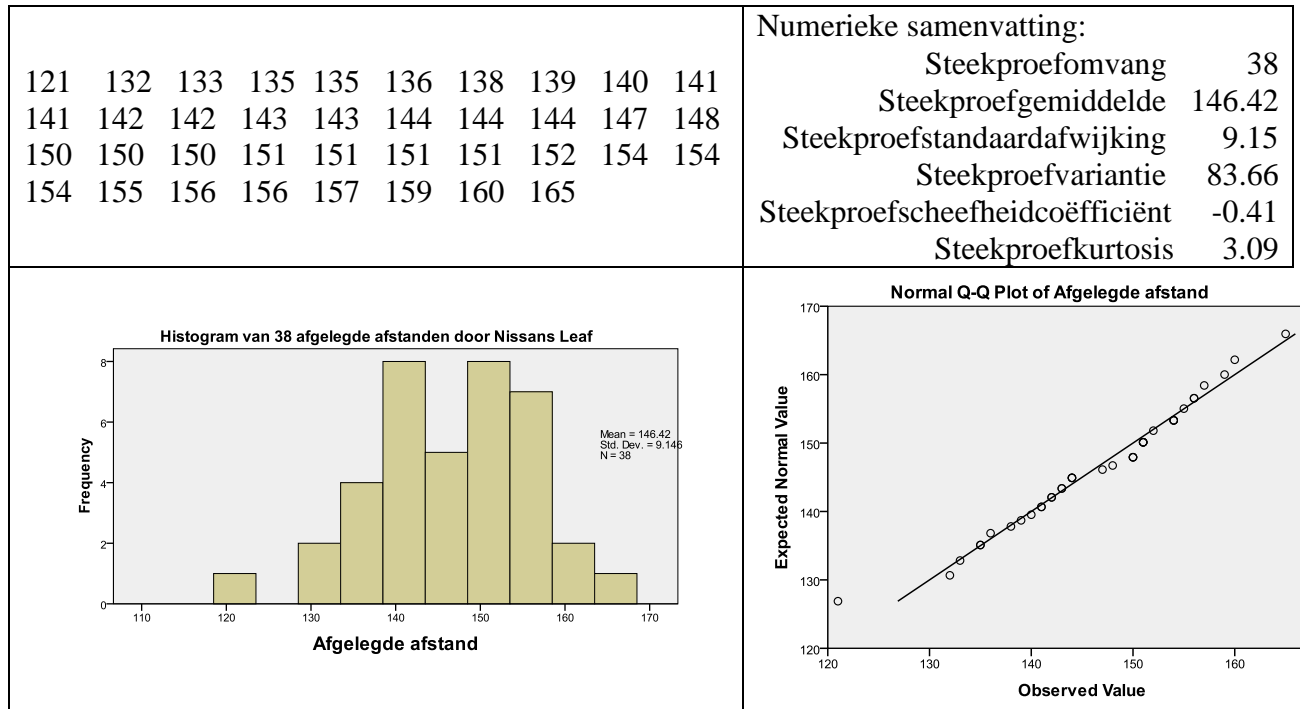
De te gebruiken tabellen komen uit het dictaat van Statistiek(TBK, INF,BIT), maar zijn hier NIET daadwerkelijk toegevoegd. De vier tabellen worden beschouwd als de standaard set van tabellen.

Opmerking 3:

Er is opzettelijk geen Poisson tabel toegevoegd. De betrokken kansen bij opgave 3 kunnen redelijkerwijs met een rekenmachine berekend worden.

Opgave 1

Een groep van 38 eigenaren van de nieuwe elektrische auto Nissan Leaf is bereid gevonden mee te doen aan een test om de werkelijke actieradius van hun auto bij een volle accu te bepalen (volgens de producent ca. 160 km). Bij een voor hen normaal gebruik rapporteerden zij de volgende afstanden (gerangschikt van klein naar groot). Tevens zijn een numerieke samenvatting en een tweetal grafieken bepaald. De vraag is of voor de afgelegde afstanden de normale verdeling als model gebruikt kan worden. We gaan er vanuit dat deze waarnemingen opgevat kunnen worden als een aselecte steekproef uit de betreffende populatie.



Bepaal volgens de box plot methode of er (potentiële) uitschieters zijn.

Opgave 2

X has a uniform distribution on the interval $(0, \theta)$, with unknown parameter θ .

In order to estimate θ we have a random sample X_1, \dots, X_n of X .

- a. Show that $T_1 = \max(X_1, \dots, X_n)$ is the *mle* of θ .
- b. Is T_1 an unbiased estimator of θ ?
- c. Is T_1 a consistent estimator of θ ?

Since $E(X) = \frac{\theta}{2}$, intuitively $T_2 = 2\bar{X}$ seems a good estimator of θ .

- d. Assume $n = 20$. Which of the two estimators of θ , T_1 and T_2 , is the best?

Opgave 3

Let X_1, \dots, X_{10} denote a random sample of size of a Poisson distribution with mean θ .

$$\text{So } P(X_i = x_i) = \frac{\theta^{x_i}}{x_i!} e^{-\theta} \quad (x_i = 0, 1, 2, \dots)$$

Show that the test that rejects $H_0: \theta = 0.1$ in favour of $H_1: \theta = 0.5$ if $\sum_{i=1}^{10} X_i \geq 3$ is most powerful. Determine, for this test, the significance level α and the power at $\theta = 0.5$.

Opgave 4

Beschouw onafhankelijke waarnemingen X_1, X_2, \dots, X_n die alle verdeeld volgens de dezelfde (onbekende) verdeling. De steekproefvariantie S^2 is een zuivere schatter van de onbekende variantie σ^2 ongeacht de verdeling van de waarnemingen. Toon dit aan.

Opgave 5

Geef aan hoe we een 95% betrouwbaarheidsinterval voor μ construeren, uitgaande van onafhankelijke waarnemingen X_1, X_2, \dots, X_9 die alle verdeeld zijn volgens een $N(\mu, \sigma^2)$ -verdeling met μ en σ^2 beide onbekend.

Je mag uitgaan van de bekende resultaten ten aanzien van de verdelingen van het steekproefgemiddelde \bar{X} en steekproefvariantie S^2 . Motiveer kort de te nemen stappen.

Opgave 6

Een marketing-adviesbureau is bezig met een reclamecampagne voor kleding voor meisjes van 10-12 jaar. Een belangrijk aspect bij de aankoop van kleding is wie uiteindelijk tot de koop beslist: de moeder of de dochter. Een onderzoek bij 400 aankopen wees uit dat in 243 gevallen de koopbeslissing door de moeder werd genomen. Valt bij een significantieniveau van 5% te concluderen dat de moeders in de meerderheid van alle dergelijke aankopen de eindbeslissing nemen?

- Voer een toets uit om deze vraag te beantwoorden en volg het schema van 8 stappen, zie formuleblad.
- Bepaal de kans op een fout van de tweede soort als in werkelijkheid in 60% van de gevallen de moeder beslist. Geef ook het onderscheidend vermogen voor die waarde van p .

Opgave 7

Plantenskundigen hebben maïsvariëteiten ontwikkeld die een verhoogde hoeveelheid van het essentiële aminozuur lysine bevatten. In een test naar de proteïnekwaliteit van deze maïsvariëteiten werd een experimentele groep (een aselechte steekproef) van 20 één dag oude haantjes gevoerd met een voedselpakket dat deze maïs bevatte. Een controlegroep van 20 andere haantjes kreeg een voedselpakket dat identiek was behalve dat er normale maïs in zat. Hier volgen de gewichtstoenames (in grammen) na 21 dagen.

controlegroep				experimentele groep			
380	321	366	356	361	447	401	375
283	349	402	462	434	403	393	426
356	410	329	399	406	318	467	407
350	384	316	272	427	420	477	392
345	455	360	431	430	339	410	326

Beschouw de resultaten x_1, \dots, x_{20} bij de controlegroep en y_1, \dots, y_{20} bij de experimentele groep als realisaties van de s.v.-en X_1, \dots, X_{20} en Y_1, \dots, Y_{20} .

Verder noteren we $z_i = y_i - x_i$

We hebben de volgende numerieke samenvatting berekend:

	Gemiddelde	Standaardafwijking	Aantal
x	366.30	50.80	20
y	404.75	42.73	20
z	38.45	82.68	20

- Bereken het 90%-betrouwbaarheidsinterval voor het verschil in verwachting tussen de controlegroep en de experimentele groep.
- Welke (model)veronderstellingen zijn bij onderdeel a (impliciet?) gebruikt om de formules te kunnen toepassen?

Opgave 8

Om twee koperpoetsmiddelen A en B te vergelijken heeft men 23 koperen platen op verschillende plaatsen geruime tijd blootgesteld aan weer en wind. Vervolgens zijn de platen gehalveerd en de twee helften met poetsmiddel A resp. met poetsmiddel B behandeld. Een commissie heeft alle 46 halve platen afzonderlijk (zonder de herkomst van het middel te kennen) beoordeeld door cijfers toe te kennen voor het poetsresultaat. In de onderstaande tabel zijn de oordelen van de commissie van de platen $1A$ en $1B$ t/m $23A$ en $23B$ weergegeven.

plaat	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
A	9	7	6	7	6	5	8	9	8	7	5	10	6	7	7	7	6	8	8	8	5	6	8
B	4	6	5	7	3	9	3	4	7	8	6	5	4	6	6	6	8	4	6	6	3	2	7

- Gaat het hier om gepaarde waarnemingen of twee onafhankelijke steekproeven?
- Kunnen we hier redelijkerwijs de normale verdeling veronderstellen voor de gepresenteerde waarnemingen?
- Onderzoek of er een systematisch verschil is tussen het oordeel over poetsmiddel A en poetsmiddel B . Pas een geschikte toets toe en volg het schema van acht stappen voor een toets, zie formuleblad. Gebruik $\alpha = 0.05$.