

Afdeling Wiskunde	Volledig tentamen Statistics
	Deeltentamen 2 Statistics
Vrije Universiteit	28 mei 2015

- **Gebruik van een (niet-grafische) rekenmachine is toegestaan.**
- Geheel tentamen: **opgaven 1,2,3,4.** Cijfer = $\frac{totaal+4}{4}$
- Deeltentamen 2: **opgaven 3,4,5.** Cijfer = $\frac{totaal+3}{3}$
- Na de correctie liggen de tentamens ter inzage bij het onderwijsbureau FEW.

SUCCES!

Opgave 1 [7 punten] **Alleen voor geheel tentamen**

- [2 punten] Met welk doel maakt men QQ -plots?
- [3 punten] Behoren de uniforme verdelingen op $[a, b]$ en $[c, d]$ met $b > a$ en $d > c$ tot dezelfde lokatie-schaalfamilie? Beargumenteer uw antwoord met een korte berekening.
- [2 punten] Geef de definitie van de likelihood-ratiostatistiek $\lambda_n(X)$ en geef een uitdrukking voor het $1 - \alpha$ (benaderend) betrouwbaarheidsgebied gebaseerd op deze toetsingsgrootheid.

Opgave 2 [14 punten] **Alleen voor geheel tentamen**

Zij $X_1, \dots, X_n$ een steekproef uit de Poisson-verdeling met parameter $\theta > 0$. De (marginale) kansmassafunctie voor X_1 is dan

$$P_\theta(X_1 = x) = \frac{\theta^x}{x!} e^{-\theta}.$$

In deze opgave beschouwen we schatters voor θ^2 .

- [5 punten] Bepaal de maximum likelihood-schatter voor θ^2 .
- [3 punten] Is deze schatter zuiver? Zo nee, bepaal dan een zuivere schatter voor θ^2 .
- [4 punten] Bepaal de a posteriori dichtheid voor θ ten opzichte van de Gamma(α, λ) a priori verdeling (zie Bijlage 1) gebaseerd op de steekproef $X_1, \dots, X_n$. Van welke verdeling is dit de dichtheid?
- [2 punten] Bepaal de Bayes-schatter voor θ^2 ten opzichte van de gegeven a priori verdeling. (U hoeft de uitdrukking voor deze schatter niet te vereenvoudigen of te herschrijven.)

Opgave 3 [8 punten]

Om de juli-temperaturen in De Bilt en Eelde te vergelijken heeft het KNMI de volgende gemiddelde juli-temperaturen verzameld:

jaar	1965	1966	1967	1968	1969	1970	1971	1972	1973	gem	var
De Bilt	13.7	14.2	15.3	14.1	18.1	15.7	17.2	17.2	17.1	15.84	2.625
Eelde	13.4	14.1	14.9	14.1	17.5	15.1	17.0	17.1	16.7	15.54	2.385
verschil	0.3	0.1	0.4	0.0	0.6	0.6	0.2	0.1	0.4	0.3	0.0475

De onderste rij geeft de verschillen weer. De waarde in de kolom gem is het gemiddelde van de rij, de waarde in de kolom var is de steekproefvariantie (S_X^2) van de rij. We nemen aan dat de waarnemingen over de verschillende jaren onafhankelijk zijn.

- a. [5 punten] Toets de nulhypothese dat er geen verschil bestaat tussen De Bilt en Eelde in verwachte temperatuur. Gebruik een onbetrouwbaarheidsdrempel van 5%. Vermeld daarbij
 - i.) een geschikt statistisch model
 - ii.) een geschikte nulhypothese
 - iii.) de toetsingsgrootheid
 - iv.) welke aannames u maakt
 - v.) de verdeling van de toetsingsgrootheid onder de (rand van de) nulhypothese
 - vi.) het kritieke gebied of de overschrijdingskans (p -waarde)
 - vii.) de conclusie van de toets.
- b. [3 punten] Stel een numeriek 95%-betrouwbaarheidsinterval op voor het verschil in verwachte temperatuur tussen De Bilt en Eelde op basis van een geschikte pivot. (Let op: *numeriek* wil zeggen: geef de getallen.)

Opgave 4 [7 punten]

Zij $X_1, \dots, X_n$ een steekproef uit de verdeling met (marginale) kansdichtheid

$$p_\theta(x) = \frac{1}{\theta} e^{-\frac{x-1}{\theta}} \mathbf{1}_{\{x \geq 1\}}$$

waar $\theta > 0$ een onbekende parameter is.

- a. [3 punten] Bepaal een voldoende en volledige statistiek.
- b. [4 punten] Bepaal een UMVZ-schatter voor θ .

Opgave 5 [12 punten] **Alleen voor deeltentamen 2**

Zij $X_1, \dots, X_n$ een steekproef uit de alternatieve verdeling met parameter $\theta \in [0, 1]$. De kansmassafunctie van de verdeling is te schrijven als

$$p_\theta(x) = \theta^x (1 - \theta)^{1-x}, \quad x \in \{0, 1\}.$$

De maximum likelihood-schatter voor θ is gelijk aan $\hat{\theta} = \bar{X}$.

- a. [3 punten] Bepaal de Fisher-informatie i_θ .
- b. [3 punten] Wat is de asymptotische verdeling van $\sqrt{n}(\hat{\theta} - \theta)$? Bepaal een benaderend $(1 - \alpha)$ -betrouwbaarheidsinterval voor θ op basis van $\hat{\theta}$ en de plug-in schatter voor i_θ .
- c. [3 punten] Bepaal de likelihood-ratiostatistiek $\lambda(X)$ gebaseerd op $X = (X_1, \dots, X_n)$ voor het toetsen van $H_0 : \theta = \theta_0$.
- d. [3 punten] Bepaal een benaderend $(1 - \alpha)$ -betrouwbaarheidsinterval voor θ op basis van de likelihood-ratiostatistiek.

Bijlage 1: Gamma-verdeling

Een stochastische grootte X die $\Gamma(\alpha, \lambda)$ -verdeeld is heeft de volgende dichtheid:

$$f_{\alpha, \lambda}(x) = \frac{1}{\Gamma(\alpha)} x^{\alpha-1} \lambda^\alpha e^{-\lambda x} \mathbf{1}_{(0, \infty)}(x)$$

met $\alpha, \lambda > 0$, waar de Gamma-functie als volgt gedefinieerd is:

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$$

met de eigenschap $\Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$. De verwachtingswaarde en variantie zijn gelijk aan $E_{\alpha, \lambda} X = \alpha / \lambda$ en $\text{var}_{\alpha, \lambda} X = \alpha / \lambda^2$.

Bijlage 2: Tabel normale verdeling

	0	1	2	3	4	5	6	7	8	9
0.0	0.5	0.504	0.508	0.512	0.516	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.591	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.648	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.67	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.695	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.719	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.758	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.791	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.834	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.877	0.879	0.881	0.883
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.898	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.937	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.975	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.983	0.9834	0.9838	0.9842	0.9846	0.985	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.989
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.992	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.994	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.996	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.997	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.998	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.999	0.999
3.1	0.999	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1	1	1	1	1	1	1	1	1	1

Tabel 1: Verdelingsfunctie van de standaard normale verdeling op het interval $[0, 4]$. De waarde in de tabel is $\Phi(x)$ voor $x = a + b/100$ met a de waarde in de eerste kolom en b het getal in de eerste rij.

Bijlage 3: Tabel t -verdeling

df	0.6	0.7	0.75	0.8	0.85	0.9	0.925	0.95	0.975	0.98	0.99	0.999
1	0.32	0.73	1	1.38	1.96	3.08	4.17	6.31	12.71	15.89	31.82	318.31
2	0.29	0.62	0.82	1.06	1.39	1.89	2.28	2.92	4.3	4.85	6.96	22.33
3	0.28	0.58	0.76	0.98	1.25	1.64	1.92	2.35	3.18	3.48	4.54	10.21
4	0.27	0.57	0.74	0.94	1.19	1.53	1.78	2.13	2.78	3	3.75	7.17
5	0.27	0.56	0.73	0.92	1.16	1.48	1.7	2.02	2.57	2.76	3.36	5.89
6	0.26	0.55	0.72	0.91	1.13	1.44	1.65	1.94	2.45	2.61	3.14	5.21
7	0.26	0.55	0.71	0.9	1.12	1.41	1.62	1.89	2.36	2.52	3	4.79
8	0.26	0.55	0.71	0.89	1.11	1.4	1.59	1.86	2.31	2.45	2.9	4.5
9	0.26	0.54	0.7	0.88	1.1	1.38	1.57	1.83	2.26	2.4	2.82	4.3
10	0.26	0.54	0.7	0.88	1.09	1.37	1.56	1.81	2.23	2.36	2.76	4.14
11	0.26	0.54	0.7	0.88	1.09	1.36	1.55	1.8	2.2	2.33	2.72	4.02
12	0.26	0.54	0.7	0.87	1.08	1.36	1.54	1.78	2.18	2.3	2.68	3.93
13	0.26	0.54	0.69	0.87	1.08	1.35	1.53	1.77	2.16	2.28	2.65	3.85
14	0.26	0.54	0.69	0.87	1.08	1.35	1.52	1.76	2.14	2.26	2.62	3.79
15	0.26	0.54	0.69	0.87	1.07	1.34	1.52	1.75	2.13	2.25	2.6	3.73
16	0.26	0.54	0.69	0.86	1.07	1.34	1.51	1.75	2.12	2.24	2.58	3.69
17	0.26	0.53	0.69	0.86	1.07	1.33	1.51	1.74	2.11	2.22	2.57	3.65
18	0.26	0.53	0.69	0.86	1.07	1.33	1.5	1.73	2.1	2.21	2.55	3.61
19	0.26	0.53	0.69	0.86	1.07	1.33	1.5	1.73	2.09	2.2	2.54	3.58
20	0.26	0.53	0.69	0.86	1.06	1.33	1.5	1.72	2.09	2.2	2.53	3.55
21	0.26	0.53	0.69	0.86	1.06	1.32	1.49	1.72	2.08	2.19	2.52	3.53
22	0.26	0.53	0.69	0.86	1.06	1.32	1.49	1.72	2.07	2.18	2.51	3.5
23	0.26	0.53	0.69	0.86	1.06	1.32	1.49	1.71	2.07	2.18	2.5	3.48
24	0.26	0.53	0.68	0.86	1.06	1.32	1.49	1.71	2.06	2.17	2.49	3.47
25	0.26	0.53	0.68	0.86	1.06	1.32	1.49	1.71	2.06	2.17	2.49	3.45
26	0.26	0.53	0.68	0.86	1.06	1.31	1.48	1.71	2.06	2.16	2.48	3.43
27	0.26	0.53	0.68	0.86	1.06	1.31	1.48	1.7	2.05	2.16	2.47	3.42
28	0.26	0.53	0.68	0.85	1.06	1.31	1.48	1.7	2.05	2.15	2.47	3.41
29	0.26	0.53	0.68	0.85	1.06	1.31	1.48	1.7	2.05	2.15	2.46	3.4
30	0.26	0.53	0.68	0.85	1.05	1.31	1.48	1.7	2.04	2.15	2.46	3.39
31	0.26	0.53	0.68	0.85	1.05	1.31	1.48	1.7	2.04	2.14	2.45	3.37
32	0.26	0.53	0.68	0.85	1.05	1.31	1.47	1.69	2.04	2.14	2.45	3.37
33	0.26	0.53	0.68	0.85	1.05	1.31	1.47	1.69	2.03	2.14	2.44	3.36
34	0.26	0.53	0.68	0.85	1.05	1.31	1.47	1.69	2.03	2.14	2.44	3.35
35	0.26	0.53	0.68	0.85	1.05	1.31	1.47	1.69	2.03	2.13	2.44	3.34
36	0.26	0.53	0.68	0.85	1.05	1.31	1.47	1.69	2.03	2.13	2.43	3.33
37	0.26	0.53	0.68	0.85	1.05	1.3	1.47	1.69	2.03	2.13	2.43	3.33
38	0.26	0.53	0.68	0.85	1.05	1.3	1.47	1.69	2.02	2.13	2.43	3.32
39	0.26	0.53	0.68	0.85	1.05	1.3	1.47	1.68	2.02	2.12	2.43	3.31
40	0.26	0.53	0.68	0.85	1.05	1.3	1.47	1.68	2.02	2.12	2.42	3.31

Tabel 2: (Beneden-) Kwantielen van de t -verdelingen met 1 tot 40 vrijheidsgraden.