

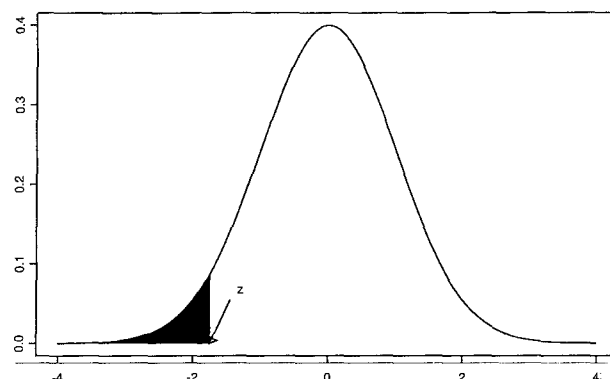


Tentamen Toegepaste Statistiek
voor studenten Informatiekunde en Kunstmatige Intelligentie
vrije Universiteit, Faculteit Exacte Wetenschappen
27 maart 2006, 8.45-10.45u.

1.
 - a) Onderzoek van talloze publicaties over de relatie het optreden van borstkanker en de bewegingsgewoonten van vrouwen leidt tot de conclusie dat onder vrouwen die veel aan sport doen borstkanker minder voorkomt dan onder vrouwen die weinig aan sport doen. Welk type statistische studie is dit? Beargumenteer uw antwoord.
 - b) Geef met behulp van een voorbeeld aan wat systematisch steekproeftrekken is.
 - c) Op een school wordt onderzoek gedaan naar het effect van het oefenen op de CITO toets op het uiteindelijke resultaat ervan. Per klas wordt de helft van de kinderen door de leraar aangewezen als 'oefenaars'. De andere helft oefent niet. Uiteindelijk worden de resultaten van alle scholieren bij de CITO toets bepaald en wordt gekeken of er verschil is tussen de beide groepen. Geef aan wat er bij dit onderzoek niet goed is en geef een alternatieve opzet waar u meer vertrouwen in heeft.
 - d) Geef de definitie van een kansruimte.
 - e) De tijdens het college behandelde kentallen voor locatie hebben de eigenschap dat deze locatie-equivariant zijn. Leg uit wat hiermee wordt bedoeld.
2.
 - a) Op een bepaalde plaats wordt het CO_2 -gehalte in de lucht dagelijks gemeten. Geef aan met welk type data we hier te maken hebben. Geef ook aan wat de meetschaal is.
 - b) In een call center worden continu telefoontjes voor verschillende bedrijven afgehandeld. Bij binnenkomst wordt meteen het telefoontje geklassificeerd naar bedrijf. Welk type data levert dit op en wat is de meetschaal?
3. In het kader van een onderzoek naar lengten van zestienjarigen is een steekproef van omvang 36 uit de populatie van zestienjarigen getrokken. Hieronder zijn de gegevens, uitgedrukt in centimeters, weergegeven.

163 180 170 164 170 164 185 162 167
167 159 171 161 168 178 164 168 173
175 195 158 162 171 159 164 180 182
160 158 158 181 160 165 174 169 158

- a) Geef een tabel met 4 kolommen. In de eerste de klassen, lopend van 155-159 tot en met 195-199; in de tweede de bijbehorende frequentie, de derde de relatieve frequentie en in de vierde kolom de cumulatieve relatieve frequentie.
- b) Schets op basis hiervan een histogram voor de data en beschrijf de vorm van de dataverdeling kwalitatief.



Figuur 1: *standaard normale curve corresponderend met tabel 1*

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

Tabel 2: *te verdienen punten per onderdeel*

$$\text{Cijfer} = (\text{aantal punten} + 8) / 8$$

Succes!