

|                    |                                |
|--------------------|--------------------------------|
| Afdeling Wiskunde  | Tentamen Biostatistiek 2 (MNW) |
| Vrije Universiteit | 4 maart 2011                   |

Geef een duidelijke toelichting bij je antwoorden. Je mag gebruik maken van een niet-programmeerbare rekenmachine.

Bij het maken van het tentamen heb je mogelijk de volgende kwantielen nodig:

Van de standaard normale verdeling:  $z_{0.05} = -1.64$ ,  $z_{0.025} = -1.96$ ,  $z_{0.95} = 1.64$ ,  $z_{0.975} = 1.96$ ,  
van de t-verdeling:  $t_{9,0.05} = -1.83$ ,  $t_{18,0.05} = -1.73$ ,  $t_{9,0.95} = 1.83$ ,  $t_{18,0.95} = 1.73$ .

Veel succes!

1. Veronderstel dat  $X_1, \dots, X_n$  onderling onafhankelijk en gelijk verdeeld zijn met dichtheid

$$f_\lambda(x) = 2\lambda x e^{-\lambda x^2} \quad x > 0,$$

met  $\lambda > 0$  onbekend. Bepaal de maximum likelihood schatter voor  $\lambda$ . Leg duidelijk uit hoe je aan deze schatter gekomen bent.

2. We willen onderzoeken welk percentage van de M & M's in een zakje rood is. Oftewel, als we een willekeurige M & M uit een willekeurig zakje halen, wat is dan de kans dat deze M & M rood is.

We nemen een steekproef van 30 zakjes M & M's met ieder 60 M & M's. Laat  $Y_i$  het aantal rode M & M's zijn in het  $i$ -de zakje ( $i = 1, \dots, 30$ ). Veronderstel dat  $Y_1, \dots, Y_{30}$  onderling onafhankelijk zijn en dat  $Y_i$  binomiaal verdeeld is met parameters 60 en  $\pi$ : de fractie rode M & M's in een zakje.

- (a) Bepaal de maximum likelihood schatter voor  $\pi$ .
  - (b) Bepaal de maximum likelihood schatter voor het verwachte aantal rode M & M's in een zakje.
  - (c) Bepaal de momenten-schatter voor  $\pi$ .
3. Bij de brand in het Brabantse Moerdijk zijn er mogelijk giftige stoffen vrij gekomen. Om dit te onderzoeken zijn er op verschillende plaatsen nabij de brand bodemonsters genomen. Gecontroleerd is op de aanwezigheid van aluminium. Volgens het RIVM bevatten monsters genomen van kleigrond sowieso een hoger aluminiumgehalte dan monsters van zand. Daarom wordt besloten ook monsters te nemen in kleigrond op veilige afstand van de brand. We willen nu toetsen of het aluminiumgehalte in de kleigrond nabij de brand in Moerdijk hoger is dan het aluminiumgehalte op veilige afstand. Op beide plekken zijn 10 monsters genomen. Het aluminiumgehalte in de 10 monsters nabij de brand wordt genoteerd met  $X_1, \dots, X_{10}$ , die op veilige afstand met  $Y_1, \dots, Y_{10}$ . We hebben gevonden de steekproef gemiddelden  $\bar{x} = 101.5$ ,  $\bar{y} = 99.2$  en steekproef varianties  $s_x^2 = 5.1$  en  $s_y^2 = 5.2$ .

(a) Geef een statistisch model en formuleer vervolgens bovenstaande onderzoeksvraag als een toetsingsprobleem. Beschrijf de standaard toets. Geef de toetsingsgrootte en het kritieke gebied. Gebruik een onbetrouwbaarheidsdrempel van 5%.

(b) Verwerp je de nul hypothese? Wat is je conclusie?

Een andere onderzoeker redeneert als volgt. Het aluminium heeft tijd nodig om met het regenwater naar dieper gelegen grondlagen te spoelen. Hij wil daarom onderzoeken of het aluminiumgehalte in de bovenste laag hoger is dan in de onderste laag; een verhoging in de bovenste laag kan zijn veroorzaakt door de brand. Hij bepaalt voor elk monster dat gedaan was nabij de brand in Moerdijk het aluminiumgehalte in de bovenste en de onderste laag.

(c) Formuleer de bovenstaande onderzoeksvraag als een toetsingsprobleem. Geef de toetsingsgrootte en leg uit hoe je de p-waarde kunt bepalen.

(d) Veronderstel dat de p-value gelijk is aan 0.042. Wat is je conclusie als we een onbetrouwbaarheid  $\alpha_0$  gelijk nemen aan 0.05? Leg uit.

4. Rokers die stoppen met roken nemen vaak toe in gewicht. Bij een onderzoek naar gewichtstoename worden  $n$  personen gewogen op het moment dat zij stopten met roken en een jaar later nog eens. Alle personen rookten gedurende het jaar niet. Men is geïnteresseerd in de onderzoeksvraag of het gewicht na het stoppen met roken toeneemt. Definieer  $X_1, \dots, X_n$  als de gewichtstoename van de  $n$  personen (waarbij  $X_i$  negatief is in geval van een afname). Veronderstel dat  $X_1, \dots, X_n$  onderling onafhankelijk en normaal verdeeld zijn met onbekende verwachting  $\mu$  en bekende standaardafwijking  $\sigma = 4$ .

(a) Formuleer de bovenstaande onderzoeksvraag als een toetsingsprobleem en geef de standaard toets. Geef de toetsingsgrootte en het kritieke gebied. Gebruik een onbetrouwbaarheidsdrempel van 5%.

(b) We willen dat als het verwachte toegenomen gewicht gelijk is aan 2 kilo, het onderscheidend vermogen ten minste 80% is. Hoe groot moeten we  $n$  minimaal kiezen om dit resultaat te bereiken? Leg uit hoe je aan dit aantal komt.

**Normering:**

|    |   |       |   |       |   |       |   |
|----|---|-------|---|-------|---|-------|---|
| 1: | 3 | 2(a): | 4 | 3(a): | 3 | 4(a): | 3 |
|    |   | 2(b): | 2 | 3(b): | 2 | 4(b): | 3 |
|    |   | 2(c): | 3 | 3(c): | 3 |       |   |
|    |   |       |   | 3(d): | 1 |       |   |

**Eindcijfer: (totaal + 3)/3**