

# Examen Statistiek en Gegevensverwerking

Fysica en Sterrenkunde 2de Bach

AJ 2021-2022, 1ste zit

---

## 1 Theorie

**1. (5pt)** - We hebben een steekproef met gekende  $x_i$  en gemeten  $y_i$  (met  $i = 1 \dots N$ ) waarbij alle  $y_i$  eenzelfde, Gaussische onzekerheid  $\sigma$  hebben. We weten ook dat  $y = f(x)$  met  $f(x) = ax$ .

(a) Uitgaande van het principe van de kleinste kwadraten, leid af wat de schatting van  $a$  en de onzekerheid hierop zijn.

(b) Laat zien dat, in dit geval, de Maximum Likelihood Methode hetzelfde resultaat geeft voor de schatting van  $a$  als de methode van de kleinste kwadraten.

(c) Leg het principe van de Maximum Likelihood Methode uit. Wat zijn de voordelen en eventuele nadelen van de methode?

**2. (3pt)** - Met behulp van de wet van foutenpropagatie, bewijs dat fractionele fouten op 2 metingen kwadratisch opgeteld kunnen worden voor zowel het product als het quotiënt van die metingen. In welke situatie is dit niet geldig? Wat zijn de onzekerheden op het product en het quotiënt als beide metingen aan dezelfde systematische fout  $S$  onderhevig zijn?

## 2 Oefeningen

**1. (4pt)** - Het gebeurt vaak dat het aantal boekingen voor een vlucht groter is dan het aantal reizigers dat aanwezig is voor de inscheping. Gedreven door scherpe concurrentie besluit de luchtvaartmaatschappij *StatAirlines* alleen inschepende reizigers te laten betaken, zonder diegenen die niet verschijnen te bestraffen. Om hiervoor te compenseren wordt er ook besloten om een groter aantal boekingen te aanvaarden dan de maximale capaciteit. Tijdens de inscheping krijgen diegenen die als eerste verschijnen een stoel toegewezen. Diegenen die later komen hebben pech als het vliegtuig vol is, maar zij ontvangen een monetaire vergoeding.

Veronderstel een vlucht met een maximale capaciteit van 300 zitplaatsen. We definiëren  $n$  als het aantal boekingen en  $S_n$  als het aantal reizigers dat aanwezig is voor inscheping. De kans om een passagier die geboekt heeft te zien bij inscheping is 90%. (De aanwezigheid van de ene reiziger is onafhankelijk van de anderen.)

(a) Wat is de waarschijnlijkheidsverdeling van  $S_n$ ? Wat zijn de verwachtingswaarde en zijn variantie?

(b) Wat is het maximale aantal boekingen  $n_{max}$  dat aanvaard kan worden zodat de kans dat 1 of meerdere reizigers de inscheping geweigerd wordt kleiner dan 5% is?

(c) Een vliegtuigticket kost 180 euro. Een overtollige klant krijgt 400 euro als vergoeding. De weringskosten voor de maatschappij bedragen 50.000 euro per vlucht. Wat is de kans met verlies te vliegen als de maatschappij 340 boekingen heeft aanvaard?

**2. (4pt)** - Een standaard extractiemethode om het loodgehalte in een bepaald bodemtype te bepalen resulteert in een gemiddelde van 86 deeltjes per miljoen (ppm). Een nieuwe, snellere extractiemethode wordt getest op 20 bodemstralen en produceert een gemiddelde van 83 ppm en een standaarddeviatie van 10 ppm.

(a) Kan men op significantieniveau 0.1 aantonen dat deze nieuwe methode gemiddeld minder extraheert dan de oude methode? Vermeld indien nodig de veronderstelling(en) die je gemaakt hebt.

(b) Aan welk type fout is deze conclusie onderhevig? Leg concreet uit.

(c) Een criticus stelt omwille van variaties in de bodem de bovenstaande test van de extractiemethode niet erg effectief is. Hoe kun je het experiment verbeteren om dit probleem op te vangen (zonder eenvoudigweg het aantal metingen te vergroten)?

**3. (4pt)** - We hebben een ruimtetelescoop met 20 gammastraling detectoren die elk in een andere richting kijken.

(a) Na 1 uur lang gegevens verzamelen hebben we volgende aantallen gemeten fotonen voor de 20 detectoren: 0,3,2,1,2,1,1,0,0,1,0,1,2,1,2,0,1,0,1,3. Toon aan dat deze verdeling consistent is met een uniforme achtergrondstraling met een gemiddelde van 1.0 fotonen per uur per detector.

(b) Neem nu aan dat er inderdaad sprake is van een uniforme achtergrondstralen met een gemiddelde van 1.0 fotonen per uur per detector. In een volgende, soortgelijke reeks gegeven zoals hierboven gegeven zien we dat 1 of meer van de detectoren meer dan 3 fotonen telt. Interpreteer deze waarneming.

(c) Veronderstel nu dat ze dankzij een onafhankelijke waarneming verwachten een discrete bron van gammastraling te zien met 1 specifieke detector. Kies een gewenst significantie niveau, en bereken hoeveel fotonen we zullen moeten meten met die detector om de conclusie te kunnen trekken dat we de veronderstelde bron van gammastraling hebben waargenomen.