

Examenvragen: Theorie

Gegevens kunnen gemeten worden op verschillende types schalen: welke ? Indien je een dataset opmeet met de persoonlijke gegevens van de inwoners van Brussel, heb je data zoals: naam, geboortejaar, gewicht, lengte, geslacht,...
Geef de schaal van deze gegevens.

Welke kengetallen ken je om de spreiding van de gegevens in een steekproef weer te geven. Definieer ze. Hoe maak je een boxplot ?

Welke kengetallen ken je om de centrale tendens van de gegevens in een steekproef weer te geven. Definieer ze.

Welke kengetallen ken je om de positie van een datapunt in een steekproef weer te geven. Definieer ze.

Indien je over een steekproef beschikt, wat is dan de relatieve en de cumulatieve frequentie ? Hoe kun je ze grafisch voorstellen (eventueel op meerdere manieren) ? Wat is een empirische verdelingsfunctie ?

Geef de definities van universum, gebeurtenis en het begrip kans. Wat is een voorwaardelijke kans ?

Formuleer de wet van de totale kans en de regel van Bayes. Illustreer beiden met de hulp van een figuur. Bewijs de regel van Bayes.

Formuleer de regel van Bayes en illustreer deze wet met de hulp van een voorbeeld en een figuur.

Wat betekent het indien twee of meerdere gebeurtenissen onafhankelijk zijn ? Geef één gevolg.

Wat is een stochastische variabele ? Geef tevens de definitie van een verdelingsfunctie en van een dichtheidsfunctie. Bestaat de laatste altijd ?

Wat zijn onafhankelijke stochastische variabelen ? Indien twee stochastische variabelen onafhankelijk zijn, wat weet je dan over de verwachtingswaarde van hun produkt en de variantie van hun som ?

X en Y zijn 2 stochastische variabelen. Hoe zijn we tot de definitie van het begrip covariantie van X en Y gekomen ? Geef tevens de definitie van de correlatie van X en Y . Zijn twee ongecorreleerde stochastieken altijd onafhankelijk ?

Men zegt dat een geometrisch verdeelde stochastische variabele geen geheugen heeft. Wat betekent dit ? Ken je nog een andere verdeling, die geen geheugen heeft ? Bewijs dat beide (geometrische + de andere) verdelingen geen geheugen hebben.

Geef de centrale limietstelling. Illustreer hoe je deze stelling kunt gebruiken bij het benaderen van een Poissonverdeling

Geef de centrale limietstelling. Illustreer hoe je deze stelling kunt gebruiken bij het benaderen van een binomiaalverdeling

Geef de definitie van een statistiek, van een schatter en van een zuivere schatter. Geef tevens een zuivere schatter voor de populatievariantie en bewijs zijn zuiverheid.

Geef de definitie van een statistiek, van een schatter en van een zuivere schatter. Geef tevens een zuivere schatter voor het populatiegemiddelde en geef de verdeling van deze schatter.

Je neemt een eerste steekproef van grootte n uit een normaal verdeelde populatie met gemiddelde μ_1 en variantie σ^2 . Je neemt nadien een tweede steekproef van grootte m uit een andere normaal verdeelde populatie met gemiddelde μ_2 en een variantie σ^2 . Geef de verdeling van het verschil van de 2 steekproefgemiddelden. Gebruik deze verdeling om een toets af te leiden voor het toetsen van de gelijkheid van de twee populatiegemiddeldes? Wat kan je doen indien beide populatievarianties verschillend zijn?

Leid een betrouwbaarheidsinterval af voor het populatiegemiddelde indien de populatievariantie gekend is (de populatie is normaal verdeeld). Welke factoren bepalen de lengte van dit interval? Wat verandert er aan de uitdrukking voor dit interval indien de populatievariantie niet gekend is?

Leid een betrouwbaarheidsinterval af voor de populatievariantie als de populatie normaal verdeeld is.

Wat zijn de verschillende stappen bij het toetsen van hypothesen? Illustreer dit aan de hand van een toets voor het gemiddelde indien de variantie ongekend is.

Wat zijn de verschillende stappen bij het toetsen van hypothesen? Illustreer dit aan de hand van een toets voor de variantie.

Welke toetsen ken je om gemiddeldes van 2 verschillende populaties met elkaar te vergelijken? Geef een kort overzicht. Vermeld tevens wanneer je welke toets gebruikt.

Bij het toetsen van hypothesen werden begrippen als kritisch punt, significantieniveau α en een p-waarde geïntroduceerd. Wat betekenen deze begrippen? Illustreer dit aan de hand van een toets voor het vergelijken van varianties in 2 verschillende normaal verdeelde populaties.

Definieer de correlatiecoëfficiënt van 2 steekproeven en van 2 stochastische variabelen. Beschrijf de toets die je kunt uitvoeren om te testen of de correlatie significant is. Wat betekent dit?

Welke toetsen ken je om op basis van een steekproef uit je populatie na te gaan of je populatie uniform verdeeld is? Geef hun voor- en nadelen. Beschrijf de Kolmogorov-Smirnov toets.

Welke toetsen ken je om op basis van een steekproef uit je populatie na te gaan of je populatie uniform verdeeld is. Geef hun voor- en nadelen. Beschrijf de chi-kwadraat toets.

Wat is een kruistabel? Leid een toets af waarmee je de onafhankelijkheid van de kolom- en de rijvariabele kunt nagaan?

Welke types fouten heb je bij het toetsen van hypothesen. Wat is de macht van een toets ? Leid de macht van de eenzijdige t-toets af. Welke factoren bepalen deze grootte ? Schets dit.

Leid de uitdrukkingen af voor de richtingscoëfficiënt en het snijpunt met de Y-as bij lineaire regressie (kleinste kwadraten methode). Is er een verband met de covariantie ?

Aan welke voorwaarden moet er voldaan zijn om de kleinste kwadraten toe te passen. Leg uit!

Hoe kun je het lineaire kleinste kwadraten model verifiëren ? Leg uit.

Wat is de determinatiecoëfficiënt bij lineaire regressie ? Hoe kun je de significantie van je regressie toetsen ?

Leg volgende grafiek uit ? Hoe kun je aan de hand van de uitdrukkingen voor de BI's een goed experiment opzetten ? Geef de voor- en de nadelen.

In dien je een lineair model fit, krijg je soms te maken met outliers. Wat zijn dit ? Welk effect hebben ze ? Wat kun je doen ?

Bij het toetsen van hypothesen werden begrippen als kritisch punt, significantieniveau α en een p-waarde geïntroduceerd. Wat betekenen deze begrippen ? Illustreer dit aan de hand van een toets om te toetsen of het populatiegemiddelde van een normaal verdeelde populatie gelijk is aan een bepaald getal a . Maak een grafiek waarop je al de bovenstaande begrippen aanduidt.