**Overzicht en schema voor berekeningen Statistiek deel 2 Cursus 2019-2020**

In het tweede deel van Statistiek wordt veel gerekend aan **puntschatting** (één waarde) en **intervalschattingen** (**betrouwbaarheidsinterval** bij gegeven **betrouwbaarheid**) voor de **parameter van een kansverdeling** aan de hand van een **steekproef**. Bij intervalschatting kan ook bepaling van de **steekproefgrootte** aan de orde zijn om een **nauwkeurigheid** te garanderen. is de halve breedte van het betrouwbaarheidsinterval, die kan worden verkleind door een grotere steekproef te nemen. Verder moeten er bij toetsen **grenswaarden** en **kritieke gebieden** worden uitgerekend. In het boek wordt gebruik gemaakt van de tabellen voor de belangrijkste kansverdelingen. Het werken met tabellen levert allerlei beperkingen op waardoor diverse trucs moeten worden uitgehaald om waarden te kunnen berekenen die niet direct in de tabellen staan, zoals interpolatie, omzetten van negatieve naar positieve waarde, werken met z-waarden, benaderingen met normale of Poissonverdeling, die ook allemaal weer gelden met verschillende geldigheid, afhankelijk van steekproefgrootte. Dit maakt het geheel behoorlijk onoverzichtelijk.

Daarom heb ik hieronder ik een overzicht gemaakt van hoe je zo eenvoudig mogelijk in alle gevallen die je op het tentamen tegen zult komen met de GR tot resultaten kunt komen.

**LET OP: Handig om te gebruiken, maar je moet nog steeds snappen wat je doet en waarom!**

**Bernoulli/Binomiale verdeling voor parameter (succeskans, fractie)**

**Steekproefgrootte**: .

**Steekproef**: (aantal successen uit )

**Puntschatter**:

**Schattingsinterval**: met betrouwbaarheid .

**Berekening**: Clopper-Pearson, los op met GR (Let op, is kans, dus startwaarde tussen 0 en 1!):

en

**Steekproefgrootte voor nauwkeurigheid** : met . Gebruik voor een schatting, als je het niet weet is de veiligste keuze . Als de berekende groter is dan 10% van de totale populatie , dan moet de volgende formule worden toegepast: .

**Kritiek gebied voor , linkszijdig:** . is de geclaimde (extreme) waarde uit H0:

Rond naar boven of beneden af op een geheel getal zodat bij invullen < geldt i.p.v. =.

**Kritiek gebied voor , rechtszijdig:** . is de geclaimde (extreme) waarde uit H0:

Rond naar boven of beneden af op een geheel getal zodat bij invullen < geldt i.p.v. =.

**Kritiek gebied voor , tweezijdig:**  of . is de geclaimde waarde uit H0:

Rond de waarden van naar boven of beneden af op een geheel getal zodat bij invullen < geldt.

**Normale verdeling voor parameter (verwachtingswaarde), de standaarddeviatie is gegeven**

**Steekproefgrootte**: .

**Steekproef**: (meetwaarden)

**Puntschatter**:

**Schattingsinterval**: met betrouwbaarheid . Los op met de GR:

Deze berekeningen leveren hetzelfde als , dat mag ook.

**Steekproefgrootte voor nauwkeurigheid :**  met . Als de berekende groter is dan 10% van de totale populatie , pas dan de volgende formule toe:

**Kritiek gebied voor , linkszijdig:**  Los op met GR: is de geclaimde (extreme) waarde uit H0:

**Kritiek gebied voor , rechtszijdig:**  Los op met GR: is geclaimde (extreme) waarde uit H0:

**Kritiek gebied voor , tweezijdig:**  of Los op met GR: is geclaimde waarde uit H0:

**Normale verdeling voor parameter (verwachtingswaarde), de standaarddeviatie is niet gegeven**

**Steekproefgrootte**: .

**Steekproef**: (meetwaarden)

**Puntschatter**:

Bereken de steekproefstandaarddeviatie .

**Schattingsinterval**: met betrouwbaarheid . Los op met de GR:

**Steekproefgrootte voor nauwkeurigheid :**

Maak gebruik van een schatting voor en gebruik de formules van het geval dat bekend is (zie bovenaan). Als dit een waarde oplevert die kleiner is dan 30, vervang dan in de berekening door de overeenkomstige -waarde: Los die op met de GR: .

**Kritiek gebied voor , linkszijdig:**  Los op met GR: is de geclaimde (extreme) waarde uit H0:

**Kritiek gebied voor , rechtszijdig:**  Los op met GR: is geclaimde (extreme) waarde uit H0:

**Kritiek gebied voor , tweezijdig:**  of Los op met GR: is geclaimde waarde uit H0:

**Normale verdeling, parameter (standaarddeviatie)**

**Steekproefgrootte**: .

**Steekproef**: (meetwaarden)

**Puntschatter**: met

**Schattingsinterval**: met betrouwbaarheid . Los op met de GR:

**Poissonverdeling voor parameter (successen per eenheid)**

**Steekproefgrootte**: .

**Steekproef**: (aantal successen per eenheid, vaak een tijdseenheid)

**Puntschatter**:

**Schattingsinterval**: met betrouwbaarheid .

Hierin is

**Chi-kwadraatverdeling**

Wordt gebruikt om **homogeniteit** (zijn twee kansvariabelen onafhankelijk?) of **aanpassing** (is een gemeten verdeling te verklaren m.b.v. een berekende verdeling?) te toetsen, of voor een **schattingsinterval van**  (zie bij normale verdeling parameter ). is het aantal vrijheidsgraden. Voor een tabel met 1 kolom/rij van getallen: . Voor een -tabel: .

Voor toetsen met gaat het erom of de toetsingsgrootheid klein genoeg is:

Er wordt dus alleen rechtszijdig getoetst:

**Kritiek gebied voor , rechtszijdig:**  Los op met GR: