UNIVERSITEIT TWENTE
Faculteit Elektrotechniek, Wiskunde en Informatica

Tentamen Wiskundige statistiek (191530382) op dinsdag 22 januari 2013, 13.45 - 16.45 uur.

Dit tentamen bestaat uit 4 opgaven en enkele formules (zie de laatste pagina). Motiveer steeds uw antwoorden. De tabellen zijn separaat bijgevoegd. Een grafische rekenmachine is **niet** toegestaan, een wetenschappelijke (niet programmeerbare) wel.

---

1. De stochastische variabele X is exponentieel verdeeld met $EX = \theta$. $X_1, \ldots, X_n$ is een aselecte steekproef van X en we definiëren de steekproeffunctie:

$$T_1 = \frac{1}{n} \sum_{i=1}^{n} X_i$$

   a) Toon aan dat T_1 de meest aannemelijke schatter is van θ.

   b) Ga na of T_1 een zuivere schatter is.

   c) Bereken var T_1.

   d) Ga na of T_1 een consistente schatter is.

2. Het aantal auto's dat bij een kruispunt tussen 8:30 en 8:45 aankomt, wordt op 10 verschillende dagen gemeten. De metingen zijn onderling onafhankelijk:

$$78, 64, 67, 75, 62, 66, 67, 70, 75, 59$$

We gaan ervan uit dat deze verdelingsfunctie goed benaderd kan worden door een normale verdeling en gaan er dus vanuit dat het aantal auto's normaal verdeeld is met verwachting μ en variantie σ^2.

   a) Geeft een schatting voor μ met een bijbehorend 95% betrouwbaarheidsinterval. Wat betekent 95% betrouwbaarheid in relatie tot het berekende numerieke interval?

   b) Geeft een schatting voor σ^2 met een bijbehorend 95% betrouwbaarheidsinterval.

Bij nadere beschouwing lijkt een Poisson verdeling voor het aantal auto's een logische keuze. Op basis van de Poisson verdeling weten we dat $\mu = \sigma^2$. Omdat de Poisson verdeling vervelend rekent, gebruiken we weer een normale verdeling als benadering maar nu nemen we expliciet mee dat $\mu = \sigma^2$ en we parameteriseren de verdeling met één parameter $\lambda = \mu = \sigma^2$.

We gaan er dus vanaf nu nog steeds vanuit dat de stochasten normaal verdeeld zijn maar nu geparameteriseerd door $\lambda = \mu = \sigma^2$.

   c) Is het steekproefgemiddelde een zuivere schatter voor λ als gegeven is dat $\lambda = \mu = \sigma^2$?

   d) Is het gegeven dat $\mu = \sigma^2$ relevant bij het bepalen van een 95% betrouwbaarheidsinterval voor λ op basis van het steekproefgemiddelde?

   e) Bepaal een 95% betrouwbaarheidsinterval voor λ met behulp van een schatting voor λ op basis van de steekproefvariantie?

3. Een leverancier geeft een garantie van 4 jaar op zijn harde schijven. Hij ging er bij het vaststellen van die garantie vanuit dat hooguit 6% van zijn schijven binnen 4 jaar kapot zou gaan. Hij heeft nu, via zijn klanten, de levensduur van een tiental harde schijven bepaald:

$$4.6, 5.8, 5.7, 5.1, 5.5, 5.4, 4.1, 5.7, 4.3, 4.3$$

We kunnen ervan uitgaan dat de levensduur van deze harde schijven onderling onafhankelijk zijn en de leverancier gaat ervan uit dat de levensduur X normaal verdeeld is met verwachting μ en variantie σ^2.

a) Bepaal c als functie van μ en σ zodanig dat

$$P(X \leq c) = 0.06$$

De aanname van de leverancier over de levensduur van zijn computers is dus correct als $c \geq 4$.

b) Bepaal een momentenschatter voor c en bepaal daarmee een schatting voor c op basis van de waarnemingen.

c) De nulhypothese:

$$H_0: \text{ Minder dan 6\% van de schijven gaat binnen 4 jaar kapot}$$

wordt verworpen als onze momentschatter een waarde kleiner dan 4 oplevert. Kunt U intuïtief beredeneren of U verwacht dat de onbetrouwbaarheid nu kleiner dan 5% is?

Een medewerker van de UT suggereert een andere aanpak. Hij definieert een stochastische variabele Y met $Y = 0$ als $X > 4$ en $Y = 1$ als $X \leq 4$. Hij definieert $p = P(Y = 1)$. De nulhypothese kan dan geherformuleerd worden

$$H_0 : p \leq 0.05$$

d) Toets deze nulhypothese op basis van de gegeven waarnemingen zodanig dat de kans op een fout van de eerste soort hoogstens 6% is.

e) Bepaal een naar boven begrensd betrouwbaarheidsinterval voor p met betrouwbaarheid 0.95.

4. Er is een nieuw operating system Windows 8 uitgekomen waarvan de leverancier beweert dat het zorgt voor minder energiegebruik van de computer vergeleken met Windows 7. Om dit te toetsen zijn we bij een bedrijf die Windows 7 geimplementeerd heeft langsgegaan en hebben van 15 computers hun energieverbruik gemeten (de waarnemingen zijn reeds gerangschikt van klein naar groot):

$$10.3168, 11.8154, 12.1865, 12.2109, 12.5303, 13.0744, 13.2932, 13.3651,$$
$$13.4719, 13.6714, 14.4051, 14.4174, 14.8326, 15.5012, 16.3884 \qquad \text{(kWh)}$$

Een dag later zijn we bij een ander bedrijf langsgegaan die Windows 8 gebruikt en daar hebben we ook van 15 computers het energieverbruik gemeten:

$$9.5718, 10.2727, 10.2885, 11.1351, 11.6702, 11.7586, 11.8978, 11.9699,$$
$$12.1547, 12.3129, 12.3192, 13.2554, 13.3703, 14.1865, 14.2185 \qquad \text{(kWh)}$$

Een eerste test liet zien dat de data niet goed benaderd kunnen worden door een normale verdeling.

Formuleer een toets om na te gaan of de aanname dat Windows 7 computers meer energie verbruiken dan Windows 8 computers met onbetrouwbaarheidsdrempel $\alpha = 0.10$ verworpen moet worden en trek je conclusie aan de hand van de meetgegevens

---

Enige formules bij het tentamen Wiskundige Statistiek

Toets van Wilcoxon

$$EW = \tfrac{1}{2}m(m+n+1), \qquad \operatorname{var} W = \tfrac{1}{12}mn(m+n+1),$$

Exponentiële verdeling ($x \geq 0$):

$$f(x) = \lambda e^{-\lambda x}, \quad E(X) = \frac{1}{\lambda}, \quad \operatorname{var}(X) = \frac{1}{\lambda^2} \text{ en } F(x) = 1 - e^{-\lambda x}$$

Aannemelijkheidsquotiënt:

$$\Lambda = \frac{\sup\limits_{\theta \in \Theta_0} \prod\limits_{i=1}^{n} g(x_i, \theta)}{\sup\limits_{\theta \in \Theta} \prod\limits_{i=1}^{n} g(x_i, \theta)}$$

Tabellen voor de binomiale, Poisson, normale, student, chikwadraat en F-verdeling zijn bijgevoegd.

Voor de vraagstukken kunnen de volgende aantallen punten worden behaald:

Vraagstuk	1a.	3 punten	Vraagstuk	1b	1 punten	Vraagstuk	1c.	2 punten
Vraagstuk	1d.	2 punten	Vraagstuk	2a.	3 punten	Vraagstuk	2b.	2 punten
Vraagstuk	2c.	1 punten	Vraagstuk	2d.	2 punten	Vraagstuk	2e.	3 punten
Vraagstuk	3a.	2 punten	Vraagstuk	3b.	2 punten	Vraagstuk	3c.	2 punten
Vraagstuk	3d.	4 punten	Vraagstuk	3e	3 punten	Vraagstuk	4.	6 punten