Tilastollinen päättely I 12. 5. 2017
Kurssikoe 2h 30min

1. Olkoon $X_1, X_2, \ldots, X_n \sim \text{Poisson}(\mu)$ riippumattomia, $\mu > 0$. Poissonin jakaumaa noudattavalla satunnaismuuttujalla on pistetodennäköisyysfunktio

$$f(x;\mu) = e^{-\mu}\frac{\mu^x}{x!}, \qquad x = 0, 1, 2, \ldots$$

Johda huolellisesti parametrin μ suurimman uskottavuuden estimaattori $\hat{\mu}$.

2. Olkoon $X_1, X_2, \ldots, X_n$ riippumattomia $\mathsf{N}(\mu, \sigma^2)$-jakautuneita satunnaismuuttujia, missä $\mu, \sigma^2 \in \mathbb{R}$ ovat tuntemattomia parametreja ja $\sigma^2 > 0$. Merkitään $\overline{X} = n^{-1}\sum_{i=1}^{n} X_i$.

   (i) Tarkastellaan parametrin μ (kiinnostusparametri) luottamusväliä ja satunnaismuuttujia

   $$Y_1 = \frac{\overline{X} - \mu}{\sigma^2/\sqrt{n}}$$

   ja

   $$Y_2 = \frac{\overline{X} - \mu}{S/\sqrt{n}},$$

   missä S on otoskeskihajonta. Luottamusväli voidaan muodostaa helposti käyttäen sopivaa saranasuuretta. Onko Y_1 saranasuure? Entä Y_2? Perustele.

   (ii) Oletetaan nyt, että $\sigma^2 = 1$ on tunnettu. Esitä kaksisuuntainen 95%-luottamusväli parametrille μ. Kuinka suuri otoskoko olisi valittava, että luottamusvälin leveydeksi tulisi noin 1? Entä jos valittaisiin vastaava 90%-luottamusväli?

3. Viesti lähetetään koodattuina biteiksi, eli se koostuu nollista ja ykkösistä. Viestistä $1/10$ on ykkösiä, ja $9/10$ nollia. Lähetysyhteys on kuitenkin huono, joten viestiä välitettäessä $1/5$ nollista muuttuu ykkösiksi, ja $1/3$ ykkösistä nolliksi.

   (i) Laske todennäköisyys, että viestiä vastaanotettaessa yksittäinen (satunnaisesti valittu) merkki on ykkönen.

   (ii) Viestiä vastaanotettaessa merkki on ykkönen. Laske todennäköisyys, että se oli ykkönen myös alkuperäisessä viestissä.

4. Havainnot ovat luvut $y_1, \ldots, y_n$. Oletamme, että nämä ovat satunnaismuuttujien $Y_1, \ldots, Y_n$ havaittuja arvoja ja että $Y_1, \ldots, Y_n$ ovat riippumattomia ja noudattavat jakaumaa $\mathsf{N}(\mu, \sigma^2)$, jossa sekä μ että σ^2 ovat tuntemattomia. Haluamme testata hypoteesia $\mu \geq \mu_0$, kun vastahypoteesina on $H_1: \mu < \mu_0$. Testi perustuu testisuureeseen

$$t = \frac{\overline{y} - \mu_0}{s/\sqrt{n}}.$$

   (i) Selitä, kuinka tämän testin p-arvo määritellään ja selitä, kuinka sen saisi laskettua esimerkiksi taulukoiden tai tietokoneiden avulla. (Voit ajatella, että nollahypoteesi on $\mu = \mu_0$).

   (ii) Testi pyydetään suorittamaan merkitsevyystasolla $\alpha = 0.05$. Mitä tarkoittaa hylkäämisvirhe (omin sanoin) ja mitä voidaan sanoa hylkäämisvirheen todennäköisyydestä tässä testitilanteessa?

   (iii) Oletetaan, että testin p-arvoksi tuli $p = 0.09$. Hylätäänkö nollahypoteesi H_0 vai jääkö se voimaan, kun merkitsevyystaso on $\alpha = 0.05$?

Standardinormaalijakauman yläkvantiileja:

α	0.1	0.05	0.025	0.01	0.005
z_α	1.28	1.64	1.96	2.33	2.58