Todennäköisyyslaskenta IIa, syys-lokakuu 2019 / Hytönen
Kurssitentti 25.10.2019 – **Ratkaisut**

Sallitut omat tarvikkeet:

- Kirjoitusvälineet.
- Ylioppilaskirjoituksissa hyväksyttävä laskin.
- Yksi käsinkirjoitettu (saa olla molemmin puolin) korkeintaan A4-kokoinen ”lunttilappu”.

*

Yritä ratkaista kaikki tehtävät. Muista perustella ratkaisusi. Hyvä ajatus on kirjoittaa auki käyttämäsi määritelmät, vaikka sitä ei erikseen pyydettäisi. Kirjoita jokaiseen arvosteltavaan paperiin nimesi, opiskelijanumerosi, kurssin nimi ja kokeen päivämäärä. Rauhoitu, keskity, menesty.

*

1. Merkitse kustakin alla olevasta funktiosta, onko se

   D = diskreetin satunnaismuuttujan kertymäfunktio;

   J = jatkuvan satunnaismuuttujan kertymäfunktio;

   K = kertymäfunktio, mutta ei kumpikaan edellisistä;

   E = ei lainkaan kertymäfunktio.

   (Huom! Sama vastaus voi esiintyä useasti; jokin vastaus ei ehkä lainkaan.)

$$F_0(x) = \begin{cases} e^x, & x < -1, \\ 1, & x \geq -1, \end{cases} \qquad F_1(x) = \begin{cases} 0, & x \leq 0, \\ 1 - e^{-x}, & x > 0, \end{cases}$$

$$F_2(x) = \begin{cases} 0, & x < 0, \\ x, & x \in [0, 1], \\ 1, & x > 1, \end{cases} \qquad F_3(x) = \begin{cases} 0, & x < -7, \\ \frac{1}{2}, & x \in [-7, -5), \\ 1, & x \geq -5. \end{cases}$$

   Voisiko jokin (mikä?) yllä olevista kertymäfunktioista järkevästi mallintaa

   (a) satunnaisen puhelinsoiton kestoa?

   (b) kolikon heiton tulosta, jos kruunalle ja klaavalle valitaan sopivat lukuarvot?

   (Tässä tehtävässä on siis $4 + 2 = 6$ lyhyttä kysymystä, kukin à 1 piste.)

   Ratkaisu. Kertymäfunktio on kasvava, oikealta jatkuva ($F(x+) = F(x)$) ja toteuttaa normalisoinnin $F(-\infty) = 0$, $F(+\infty) = 1$. Kaikilla F_i on nämä ominaisuudet.

   F_0 on **K**. Hyppy kohdassa $x = -1$ (ei ole J), saa kaikki arvot välillä $(0, e^{-1})$ (ei ole D).

   F_1 on **J**: jatkuva kaikkialla ja paloittain jatkuvasti derivoituva väleillä $(-\infty, 0)$ ja $(0, \infty)$. Voidaan myös tunnistaa jatkuvan eksponenttijakauman Exp(1) kertymäfunktioksi.

   F_2 on **J**: jatkuva kaikkialla ja paloittain jatkuvasti derivoituva väleillä $(-\infty, 0)$, $(0, 1)$ ja $(1, \infty)$. Voidaan myös tunnistaa jatkuvan tasajakauman U(0, 1) kertymäfunktioksi.

   F_3 on **D**: paloittain vakio, kaksi (erityisesti numeroituva määrä!) hyppykohtaa.

   (a) F_1, eksponenttijakauma sopii tähän.

   (b) F_3 sopii, kun kruunalle ja klaavalle valitaan lukuarvot -7 ja -5.

   Pisteytyksestä: Kustakin oikeasta vastauksesta 1 piste.

2. Edellisessä tehtävässä on kutakin funktioita F_0, F_1, F_2, F_3 koskien 4 eri vastausvaihtoehtoa (nimittäin D, J, K, E), ja kysymyksiä (a) ja (b) koskien 4+1 eri vastausvaihtoehtoa (nimittäin mainitut 4 funktiota sekä vaihtoehto ”ei mikään niistä”). Eräs opiskelija ei osaa lainkaan vastata näihin kysymyksiin vaan valitsee umpimähkään jonkin vastausvaihtoehdon jokaiseen kohtaan. Laske tällä tavalla vastaavan opiskelijan edellisestä tehtävästä saamien pisteiden

   (a) odotusarvo,

   (b) varianssi.

   Ratkaisu. Olkoon X_i pistemäärä funktiota F_i koskevasta kysymyksestä, $i = 0, 1, 2, 3$. Olkoon Y_j pistemäärä kysymyksestä (j), missä $j = a, b$. Olkoon $X = \sum_{i=0}^{3} X_i$ ja $Y = Y_a + Y_b$. Koko tehtävän pistemäärä on $X + Y$.

   Funktiokysymyksissä on 4 vaihtoehtoa, joista 1 antaa yhden pisteen ja 3 antaa 0 pistettä. Siis $P(X_i = 1) = 1/4$, $P(X_i = 0) = 3/4$ ja $EX_i = \sum_x xP(X_i = x) = 1 \cdot 1/4 + 0 \cdot 3/4 = 1/4$.

   Vastaavasti $EY_j = 1 \cdot 1/5 + 0 \cdot 4/5 = 1/5$.

   Koska $X_i \in \{0, 1\}$, niin $X_i^2 = X_i$ ja voidaan laskea

   $$\mathrm{var}X_i = E(X_i - EX_i)^2 = E(X_i^2) - (EX_i)^2 = EX_i - (EX_i)^2 = 1/4 - (1/4)^2 = 3/4^2.$$

   Vastaavasti $\mathrm{var}Y_j = 4/5^2$.

   (a) Odotusarvon lineaarisuuden perusteella

   $$E(X + Y) = EX + EY = \sum_{i=1}^{3} EX_i + \sum_{j=a,b} EY_j = 4 \cdot 1/4 + 2 \cdot 1/5 = 7/5 = 1{,}4.$$

   (b) Koska vastauksia eri kohtiin voidaan pitää riippumattomina, ja riippumattomien satunnaismuuttujien summan varianssi on varianssien summa, saadaan

   $$\mathrm{var}(X+Y) = \sum_{i=1}^{3} \mathrm{var}X_i + \sum_{j=a,b} \mathrm{var}Y_j = 4\cdot 3/4^2 + 2\cdot 4/5^2 = 3/4 + 8/25 = (75+32)/100 = 1{,}07.$$

   Vaihtoehto 1: Voidaan myös havaita, että $X_i \sim \mathrm{Bernoulli}(1/4)$ joten $X \sim \mathrm{Bin}(4, 1/4)$ ja käyttää tunnettuja tietoja $EZ = np$, $\mathrm{var}Z = np(1 - p)$, kun $Z \sim \mathrm{Bin}(n, p)$. Muuttuja $Y = Y_a + Y_b \sim \mathrm{Bin}(2, 1/5)$ käsitellään aivan vastaavasti. Lukujen $E(X + Y)$ ja $\mathrm{var}(X + Y)$ laskemiseksi tarvitaan tällöinkin odotusarvon lineaarisuutta sekä tulosta riippumattomien muuttujien summan varianssista.

   Vaihtoehto 2: Voidaan myös alkeistapaukset luetteloimalla määrittää satunnaismuuttujan $Z = X + Y$ pistetodennäköisyydet $P(Z = z)$, missä $z = 0, 1, \ldots, 6$ ja sen jälkeen laskea suoraan määritelmästä $EZ = \sum_z zP(Z = z)$ ja samoin varianssi. Tämä on erittäin työläs tapa, jossa on helppo tehdä laskuvirhe. Pari kokelasta silti onnistui tälläkin tekniikalla, ja se on sinänsä aivan oikein. Selvästi helpompi on silti edetä kuten edellä.

   Pisteytys: Kummastakin kohdasta (a) ja (b) max. 3 pistettä.

3. Olkoon $g(x) = e^{-|x|}$, $x \in \mathbb{R}$.

   (a) Määritä sellainen vakio $c \in \mathbb{R}$, että $f(x) = c \cdot g(x)$ on tiheysfunktio.

   (b) Olkoon X jatkuva satunnaismuuttuja, jonka tiheys on äskeinen f. Määritä sen momenttiemäfunktio.

   Ratkaisu:

(a / max. 2 pistettä): Tiheysfunktion on oltava ei-negatiivinen ja sen integraalin 1. Koska $g(x) > 0$, niin on erityisesti oltava $c > 0$. Lasketaan

$$\int_{-\infty}^{\infty} f(x)dx = c\int_{-\infty}^{\infty} e^{-|x|}dx = 2c\int_{0}^{\infty} e^{-x}dx = 2c\Big/_{0}^{\infty} - e^{-x} = 2c(0-(-1)) = 2c.$$

Siis $f(x)$ on tiheysfunktio jos ja vain jos

$$c = 1/2.$$

(b / max. 4 pistettä): Momenttiemäfunktio on

$$\begin{aligned} M_X(t) &= E(e^{tX}) \quad \text{(määritelmä)} \\ &= \int_{-\infty}^{\infty} e^{tx} f(x)\,\mathrm{d}x \quad \text{(muunnoksen odotusarvon laskukaava)} \\ &= \int_{-\infty}^{\infty} e^{tx}\frac{1}{2}e^{-|x|}\,\mathrm{d}x \\ &= \frac{1}{2}\Big(\int_{-\infty}^{0} e^{tx+x}dx + \int_{0}^{\infty} e^{tx-x}dx\Big) \\ &\overset{(*)}{=} \frac{1}{2}\Big(\Big/_{-\infty}^{0}\frac{e^{tx+x}}{t+1} + \Big/_{0}^{\infty}\frac{e^{tx-x}}{t-1}\Big) \\ &= \frac{1}{2}\Big(\frac{1-0}{t+1} + \frac{0-1}{t-1}\Big) = \frac{1}{2}\frac{(1-t)+(1+t)}{(1+t)(1-t)} = \frac{1}{1-t^2}. \end{aligned}$$

Yllä sijoitukset ovat äärellisiä ja sijoitukset $x = \pm\infty$ antavat nollan, kun ensimmäisessä integraalissa $t+1 > 0$ ja toisessa $t-1 < 0$, eli kaikkiaan kun $-1 < t < 1$. (Muutoin jompikumpi integraaleista on ∞, ja momenttiemäfunktio ei ole määritelty.) Momenttiemäfunktio on siis

$$M_X(t) = (1-t^2)^{-1}, \qquad t \in (-1,1).$$

Pisteytyksestä: Pienistä laskuvirheistä rangaistiin vähän, isommista käsitteellisistä virheistä enemmän. Erityisesti piti ymmärtää, että momenttiemäfunktio on ainoastaan muuttujan t funktio. Jos lopputulokseen jäi x tai X, tästä lähti useita pisteitä. Momenttiemäfunktion määrittelyväliä $t \in (-1,1)$ ei vaadittu ilmoittamaan. Myöskään sievennystä ei tarvinnut viedä juuri em. muotoon, kunhan lopputulokseksi saatiin jokin yhtäpitävä lauseke. (a)-kohdassa käsitteellisenä virheenä pidettiin, mikäli c:lle saatiin negatiivinen arvo.

4. Olkoon $X \sim N(0,1)$. Osoita, että X^2 on jatkuva satunnaismuuttuja ja määritä sen tiheys. (Standardinormaalijakauman tiheysfunktio on $\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$.)

   Ratkaisu:

   Useita tapoja, tässä eräänlainen rautalankamalli. Osoitettava, että kaikilla $a < b$ pätee $P(a \le X \le b) = \int_a^b f_{X^2}(y)\,\mathrm{d}y$ sopivalla f_{X^2} ja määritettävä tämä.

   Koska $X^2 \ge 0$ ja $P(X=0) = 0$, kun X on jatkuva, niin $P(a \le X^2 \le b) = 0$, jos $a < b \le 0$, ja $P(a \le X^2 \le b) = P(0 \le X^2 \le b)$, jos $a \le 0 < b$. Tutkitaan siis tapausta $0 < a < b$. Tällöin

$$\begin{aligned} P(a \le X^2 \le b) &= P(\sqrt{a} \le X \le \sqrt{b}) + P(-\sqrt{b} \le X \le -\sqrt{a}) \\ &= \int_{\sqrt{a}}^{\sqrt{b}} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx + \int_{-\sqrt{b}}^{-\sqrt{a}} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx \\ &= 2\int_{\sqrt{a}}^{\sqrt{b}} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx. \end{aligned}$$

Halutaan muokata tämä uuden muuttujan integraaliksi välillä $[a, b]$, joten tehdääb muuttujanvaihto $y = x^2$ eli $x = y^{1/2}$. Tällöin $dx = \frac{1}{2}y^{-1/2}dy$ ja saadaan

$$P(a \le X^2 \le b) = 2\int_a^b \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y}\frac{1}{2}y^{-1/2}dy = \int_a^b \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y}y^{-1/2}dy, \quad 0 < a < b,$$

ja kaikkiaan

$$P(a \le X^2 \le b) = \int_a^b f_{X^2}(y)dy,$$

missä

$$f_{X^2}(y) = \begin{cases} \frac{1}{\sqrt{2\pi y}}e^{-\frac{1}{2}y}, & y > 0, \\ 0, & y \le 0. \end{cases}$$

Tässä oleva lasku osoittaa samalla kertaa satunnaismuuttujan X^2 jatkuvuuden että antaa sen tiheysfunktion.

Vaihtoehto 1: Voidaan vedota johonkin muunnoksen $Y = g(X)$ tiheysfunktioita koskevaan lauseeseen, missä $g(x) = x^2$. Esimerkiksi $\mathbb{R}$ voidaan osittaa $\mathbb{R} = A_0 \cup A_1 \cup A_2$, missä $A_0 = \{0\}$, $A_1 = (-\infty, 0)$ ja $A_2 = (0, \infty)$, niin että $P(X \in A_0) = 0$ ja funktio g rajoitettuna A_i:lle on diffeomorfismi avoimelta väliltä A_i sen kuvajoukolle $B_i = g(A_i)$, kun $i > 0$. Nyt voidaan käyttää monisteen lausetta 2.13.

Vaihtoehto 2: Monisteessa on myös suoraan laskettu muunnoksen tiheysfunktio juuri kysytyssä tapauksessa $g(x) = x^2$ (esimerkki 2.10). Jos tämän muisti ulkoa (tai oli kirjoittanut lunttilapulle), tähänkin sai toki vedota.

Vaihtoehto 3: Monisteessa on myös erikseen käsitelty (jaksossa 5.3.7) normaalijakautuneen muuttujan $X \sim N(0, 1)$ neliötä, ja todettu, että se noudattaa jakaumaa $Z \sim \chi_1^2 = \mathrm{Gam}(\frac{1}{2}, \frac{1}{2})$ (khiin neliön jakauma vapausasteluvulla 1, tai vaihtoehtoisesti gammajakauma em. parametreilla). Jos vielä muisti (tai oli kirjoittanut lunttilapulle) ko. jakauman tiheysfunktion, tähänkin sai toki vedota.

Yhteenveto ratkaisutavoista: Tämä tehtävä oli periaattessa mahdollista ratkaista muistamalla ulkoa riittävästi nippelitietoa, mutta myös suoralla laskulla.

Pisteytyksestä: Kuten edellisessä, laskuvirheistä rangaistiin vähän, käsitteellisistä virheistä enemmän. Erityisesti oli ymmärrettävä, että kysytty tiheys on *funktio*, ei esim. lukuarvo (kuten vaikka odotusarvo olisi). Samoin oli ymmärrettävä, mitä satunnaismuuttujan jatkuvuus tarkoittaa (sitä, että todennäköisyydet $P(a \le Y \le b)$ saadaan integroimalla tiheysfunktiota). Sillä ei esim. ole juuri mitään tekemistä funktion $x \mapsto x^2$ jatkuvuuden kanssa.