Johdatus todennäköisyyslaskentaan, kevät 2014
Helsingin yliopisto, Matematiikan ja tilastotieteen laitos
Harjoitus 4 – Palautuspäivä ke 5.2.2014

Palauta ratkaisusi 3. kerroksen C-käytävän muovilaatikkoon. Käytä kurssitunnusta, älä opiskelijanumeroa!

Tehtäväsarja I: Ehdollinen todennäköisyys
Kertaa Tuomisen luvut 1.7–1.8.

1. Tuomisen kirjan tehtävä 1:65 (sivu 159).

2. Onko $P(A \mid B^c) = 1 - P(A \mid B)$? Perustele esim. jollakin kirjan lauseella tai vastaesimerkillä.

3. Kaverisi mielestä tapahtumien riippumattomuus tarkoittaa, että "ne ovat yhtä todennäköiset". Mitä sanot kaverillesi? Tutki onko $P(A) = P(B)$ ja ovatko A ja B riippumattomat,

(a) kun heitetään 6-sivuista ja 4-sivuista noppaa, A = "6-sivuinen noppa tuottaa ykkösen" ja B = "4-sivuinen noppa tuottaa ykkösen".

(b) kun nostetaan pakasta kaksi korttia, A = "ensimmäinen kortti on ässä" ja B = "toinen kortti on ässä".

Tehtäväsarja II: Kokonaistodennäköisyys ja Bayesin kaava
Tutustu Tuomisen lukuun 1.10.

4. Pussissa on kaksi noppaa. Toinen on tavallinen 6-sivuinen, mutta toinen on 4-sivuinen (numeroin 1,2,3,4). Pussista poimitaan umpimähkään yksi noppa, ja sitä heitetään kerran. Mitkä ovat mahdolliset heittotulokset ja mitkä ovat niiden todennäköisyydet?

5. Pussissa on kolme kolikkoa. Kolikkoa A heitettäessä kruunan todennäköisyys on 0.4, kolikolla B se on 0.5 ja kolikolla C se on 0.6. Pussista poimitaan umpimähkään yksi kolikko ja sitä heitetään kerran. Mikä on todennäköisyys saada kruuna?

6. Kuten edellä, mutta kolikkoa heitetään 10 kertaa. Merkitään D = "saatiin tasan 7 kruunaa".

(a) Laske $P(D \mid A)$.

(b) Laske $P(D \mid B)$.

(c) Laske $P(D \mid C)$.

(d) Laske $P(D)$.

Vihje: Kohdissa a–c tarkastele tilannetta toistokokeena.

7. Tilanne kuten edellä. Laske $P(A \mid D)$, $P(B \mid D)$ ja $P(C \mid D)$. (Vihje: Bayesin kaava.)

***8.** (Tuominen 1:47) Populaatiossa on 818 henkeä, joista 276 on rokotettu erästä epidemiaa vastaan. Epidemiaan sairastui 69 henkeä, joista 3 oli rokotettuja.

(a) Mikä on tn, että henkilö sairastui ehdolla, että hänet oli rokotettu?

(b) Mikä on tn, että henkilö oli rokotettu ehdolla, että hän ei sairastunut?

Tehtäväsarja III: Diskreetti satunnaismuuttuja
Tutustu Tuomisen lukuihin 2.1–2.2.

9. Noppaa heitetään kolme kertaa, ja X on satunnaismuuttuja, joka saa arvokseen suurimman tuloksista. Määritä todennäköisyys $P(X \leq 4)$. Vihje: Voit joko tarkastella kaikkia 216:ta alkeistapausta (työlästä mutta suoraviivaista), tai voit tarkastella tapahtumaa toistokokeena: onko jokainen heitto enintään nelonen?

10. Jatkoa edelliseen tehtävään. Määritä todennäköisyys $P(X \leq k)$ kaikille $k = 1, \ldots, 6$. (Olet nyt saanut selville X:n kertymäfunktion.)

11. Jatkoa edelliseen tehtävään. Määritä pistetodennäköisyys $P(X = k)$ kaikille $k = 1, \ldots, 6$. (Vihje: tarkastele kertymäfunktion arvoja peräkkäisillä kokonaisluvuilla. Mitä eroa niillä on ja miksi? Minkä tapahtuman todennäköisyyttä kertymäfunktio ilmaisee?)

12. Simuloi koetta “noppaa heitetään kolme kertaa, valitaan tuloksista suurin” esim. 1000 kertaa seuraavalla koodilla:

```
X=[]; for i=1:1000, X(i)=max(noppa(3)); end
```

Piirrä tuloksista pylväsdiagrammi ja vertaa edellisessä tehtävässä laskemiisi todennäköisyyksiin.

***13.** Herra K lähettää ystävälleen viisi kirjaa, joiden arvot ovat 10, 10, 20, 20 ja 30 euroa. Postipaketin tn kadota matkalla on 0.1. Herra K aprikoi, lähettääkö kirjat yhtenä vai viitenä eri pakettina. Selvitä kumpi on parempi lähetystapa ja miksi, jos tarkoituksena on saada

(a) mahdollisimman suuri todennäköisyys, että ystävä saa ainakin yhden kirjan,

(b) mahdollisimman suuri todennäköisyys, että kaikki kirjat tulevat perille,

(c) mahdollisimman pieni odotusarvo tappiolle, joka syntyy kadonneista kirjoista (ts. niiden rahallisesta arvosta). (Vihje: Odotusarvon lineaarisuus helpottaa laskua hyvin paljon.)

14. (Tuominen 2:9) Henkilöt A ja B heittävät vuorotellen noppaa siten, että A aloittaa. Pelin voittaa se, joka saa ensimmäisen kuutosen. Määritä A:n ja B:n voittotodennäköisyydet.

(Vihje: Nimitetään yhdeksi *vuoropariksi* sitä, että A ja B heittävät kumpikin kerran. Jos peliä on jatkunut jo $n - 1$ vuoroparia, mikä on todennäköisyys, että peli ratkeaa n:nnellä vuoroparilla? Jos näin käy, mikä on todennäköisyys, että voittaja on A? Kyseessä on ehdollinen todennäköisyys. Onko siitä apua tehtävän ratkaisemisessa?)

15. Jatkoa edelliseen tehtävään. Jos A voittaa, hän saa 100 euroa. Jos A häviää, hän maksaa x euroa.

(a) Mikä on A:n saaman rahan odotusarvo (maksaminen lasketaan negatiiviseksi)?

(b) Miten suuri x saa olla, jotta peli on A:lle kannattava siinä mielessä, että saadun rahan odotusarvo on positiivinen?

Tehtäväsarja IV: Satunnaislukugeneraattori (Ylimääräisiä tehtäviä)

Nämä ovat ylimääräisiä, täysin vapaaehtoisia tehtäviä satunnaislukugeneraattoreista kiinnostuneille. Niistä ei saa pisteitä.

Y1. Niin sanotussa lineaarikongruenssigeneraattorissa uusi "satunnaisluku" tuotetaan edellisestä aina samalla (deterministisellä) säännöllä

$$x_{n+1} = (ax_n + c) \bmod m$$

jossa mod tarkoittaa jakojäännösoperaatiota (jaetaan m:llä ja otetaan jakojäännös). Kurssin Matlab-hakemistosta löydät funktion `lcg`, jolla saat haluamasi pituisen jonon tällaisia lukuja, kun alkuarvo $x_1 = 1$. Kiinteiksi parametreiksi on valittu $a = 41$, $c = 7$ ja $m = 100$. Koska $m = 100$, niin tuotetut luvut ovat aina kokonaislukuja välillä $0, \ldots, 99$.

(a) Tuota generaattorilla sata lukua kutsulla `x = lcg(100)`. Mitä mieltä olet lukujen satunnaisuudesta? Voit katsoa itse lukujonoa ja/tai tehdä niistä pylväsdiagrammin (`pylvas(x)`). Mitä näet pylväsdiagrammista?

(b) Tuota 200 peräkkäistä lukua. Mitä tapahtuu 101:nnellä luvulla `x(101)`? Mitä tapahtuu sen jälkeen ja miksi?

(c) Oletetaan, että lukuja on tuotettu tällä generaattorilla hyvin pitkä jono. Tunnet säännön, jolla ne on tuotettu (ylempänä oleva yhtälö), mutta et tiedä missä kohtaa jonoa ollaan. Mikä on $P(x_i = 83)$?

(d) Kuten b-kohdassa, et tiedä missä kohtaa jonoa ollaan, mutta tiedät, että viimeksi on saatu luku $x_i = 36$. Mitä tiedät luvusta x_{i+1}? Mikä on ehdollinen todennäköisyys $P(x_{i+1} = 83 \mid x_i = 36)$?

Y2. Jatkoa edelliseen tehtävään. Generaattorin tuottamista kokonaisluvuista (jotka ovat välillä $0, \ldots, 99$) otetaan jokaisesta vain ensimmäinen numero (joka siis on välillä $0, \ldots, 9$). Tämän voi tehdä operaatiolla `y = floor(x/10)`. Tutki näin syntynyttä sadan luvun jonoa.

(a) Vaikuttaako jono `y` satunnaiselta? Missä mielessä? Tutki esim. numerojonoa sellaisenaan tai pylväsdiagrammia, joka kertoo montako kertaa mikäkin numero esiintyy jonossa.

(b) Ajatellaan, että jono `x` kuvaa generaattorin “sisäistä tilaa”, jota et tunne. Tunnet vain jonon `y`. Tiedät, että viimeksi on saatu numero $y_i = 3$. Mitä tiedät seuraavasta luvusta y_{i+1}? (Vihje: Komennolla `i = find(y==3)` löydät ne indeksit, joiden kohdalla `y`-vektorissa on kolmonen. Montako kertaa kolmonen esiintyi? Nyt komennolla `y(i+1)` voit katsoa, mitkä luvut tulivat näiden kolmosten jälkeen. Jos tiedät että viimeksi on saatu kolmonen, pystytkö pelkästään sen perusteella ennustamaan, mikä numero seuraavaksi saadaan?

(c) Jos tiedät kaksi viimeksi saatua numeroa, pystytkö sen perusteella ennustamaan seuraavan numeron?