Todennäköisyyslaskenta I, kesä 2017
Helsingin yliopisto/Avoin Yliopisto
Harjoitus 2, ratkaisuehdotuksia

1. Eräässä väestössä on $N = 2\,000$ ihmistä, joista 1000 kannattaa puoluetta A, 600 kannattaa puoluetta B ja 400 kannattaa puoluetta C. Poimitaan väestöstä umpimähkään ilman takaisinpanoa 3 henkilöä. Millä todennäköisyydellä otoksessa on pelkästään puolueen A kannattajia?

Ratkaisuehdotus:

Otanta ilman takaisinpanoa. Meitä kiinnostaa vain se, että puoluetta A kannattaa 1000 henkilöä kahdesta tuhannesta, ei se miten loput henkilöt ovat jakautuneet B:n ja C:n kesken (molemmat ovat "ei-A").

Tapahtumalle, että $n = 3$ henkilön otoksessa on tasan $k = 3$ A-kannattajaa ja $n - k = 0$ ei-A-kannattajaa, saadaan tn

$$\begin{aligned}\frac{\binom{1000}{3}\binom{1000}{0}}{\binom{2000}{3}} &= \frac{(1000)_3/(3!)}{(2000)_3/(3!)} = \frac{(1000)_3}{(2000)_3} \\ &= \frac{1000 \cdot 999 \cdot 998}{2000 \cdot 1999 \cdot 1998} = \frac{997002000}{7988004000} \approx 0.1248.\end{aligned}$$

Toinen tapa on laskea vaiheittain ehdollisilla todennäköisyyksillä ja kertolaskusäännöllä. Ensin poimittu on A-kannattaja tn:llä $\frac{1000}{2000}$. *Jos* ensin poimittu oli A-kannattaja, seuraava on A-kannattaja tn:llä $\frac{999}{1999}$. Sen jälkeen kolmas on A-kannattaja tn:llä $\frac{998}{1998}$. Todennäköisyydeksi saadaan siis $\frac{1000}{2000} \cdot \frac{999}{1999} \cdot \frac{998}{1998} = 0.1248$, joka on sama kuin toisella laskutavalla.

2. Jatkoa edelliseen tehtävään. Laske sama siten, että otos muodostetaan takaisinpanolla.

Ratkaisuehdotus:

Otanta takaisinpanolla eli toistokoe. Kukin poimittu henkilö on A-kannattaja tn:llä $1000/2000 = 1/2$, joten kysytty tn on $(1/2)^3 = 1/8 = 0.125$.

3. *Painotettu noppa.* Erästä noppaa on painotettu siten, että tulosten $1, 2, 3, 4, 5, 6$ todennäköisyydet ovat $0.1, 0.1, 0.1, 0.2, 0.2$ ja 0.3 (tässä järjestyksessä). Olkoon $A = \{2, 4, 6\}$ parillisten tulosten joukko ja $B = \{4, 5, 6\}$ "suurten tulosten joukko". Laske nyt todennäköisyydet $P(A)$, $P(B)$, $P(A \cap B)$ ja $P(A \cup B)$.

Ratkaisuehdotus:

Yksittäisten alkeistapausten todennäköisyydet on annettu tehtävässä, eli esim. $P(\{2\}) = 0.1$.

Joukko $A = \{2, 4, 6\}$ voidaan ajatella yhdisteeksi kolmesta alkeistapauksesta — tai joukko-opillisesti tarkkaan ottaen yhdisteeksi kolmesta *yhden alkion joukosta* $\{2\}$, $\{4\}$ ja $\{6\}$ — jotka ovat erilliset ja joiden todennäköisyydet on annettu. Äärellisen additiivisuuden (lause 1.3.3) nojalla:

$$P(A) = P(\{2, 4, 6\}) = P(\{2\}) + P(\{4\}) + P(\{6\}) = 0.1 + 0.2 + 0.3 = 0.6.$$

Samaan tapaan

$$\begin{aligned}P(B) &= P(\{4, 5, 6\}) = 0.2 + 0.2 + 0.3 = 0.7, \\ P(A \cap B) &= P(\{4, 6\}) = 0.2 + 0.3 = 0.5, \\ P(A \cup B) &= P(\{2, 4, 5, 6\}) = 0.1 + 0.2 + 0.2 + 0.3 = 0.8.\end{aligned}$$

Yhdisteen todennäköisyys olisi voitu laskea myös yhteenlaskukaavalla (lause 1.3.4)

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.6 + 0.7 - 0.5 = 0.8.$$

4. Oletetaan, että $P(A) = 0.3$ ja $P(B) = 0.4$. Paljonko luku $P(A \cap B)$ on **(a)** enintään, **(b)** vähintään?

Ratkaisuehdotus:

(a) Leikkaus $A \cap B$ voi suurimmillaan olla koko joukko A, nimittäin jos $A \subset B$ (piirrä kuva). Tällöin $P(A \cap B) = P(A) = 0.3$.

(b) Leikkaus on pienimmillään tyhjä, nimittäin jos joukot A ja B ovat erilliset. Tällöin $P(A \cap B) = P(\varnothing) = 0$ lauseen 1.3.2(i) nojalla.

5. Pussissa on kaksi noppaa: yksi tavallinen kuusisivuinen (sivujen numerot $1, \ldots, 6$) ja yksi nelisivuinen (sivujen numerot $1, \ldots, 4$). Kummassakin nopassa sivut ovat keskenään yhtä todennäköiset. Pussista poimitaan umpimähkään yksi noppa, ja sitä heitetään kerran. Mitkä ovat heittotulosten $1, \ldots, 6$ todennäköisyydet? (Käytä kokonaistodennäköisyyden kaavaa.)

Ratkaisuehdotus:

Merkitään K = "poimitaan kuusisivuinen noppa" (sen komplementti on K^c = "poimitaan nelisivuinen noppa"). Merkitään tapahtumalla A_i sitä, että heitetty noppa näyttää lukua i. Luvuille $i = 1, 2, 3, 4$ on

$$P(A_i) = P(K)P(A_i \mid K) + P(K^c)P(A_i \mid K^c) = \frac{1}{2} \cdot \frac{1}{6} + \frac{1}{2} \cdot \frac{1}{4} = \frac{5}{24}.$$

Luvuille $i = 5, 6$ on

$$P(A_i) = P(K)P(A_i \mid K) + P(K^c)P(A_i \mid K^c) = \frac{1}{2} \cdot \frac{1}{6} + \frac{1}{2} \cdot 0 = \frac{1}{12},$$

huom. nelisivuinen noppa ei koskaan näytä lukua 5 tai 6, joten näiden lukujen ehdolliset tn:t ovat nollia.

6. Eräässä kunnassa on 1000 asukasta. Kunta jakautuu neljään alueeseen A, B, C ja D, joiden asukasluvut ovat 100, 100, 300 ja 500. Kamreeri K valitsee ensin umpimähkään (symmetrisesti) yhden alueista ja sitten kyseiseltä alueelta umpimähkään (symmetrisesti) yhden asukkaan.

(a) Mikko asuu alueella A. Mikä on hänen todennäköisyytensä tulla valituksi?

(b) Ville asuu alueella D. Mikä on hänen todennäköisyytensä tulla valituksi?

Ratkaisuehdotus:

Merkitään tapahtumilla A, B, C, D samannimisten alueiden valintaa, jolloin umpimähkäisessä eli symmetrisessä valinnassa on $P(A) = P(B) = P(C) = P(D) = 1/4$.

(a) Merkitään Mikon valintaa tapahtumalla M. Mikko voi tulla valituksi vain, jos valittiin alue A, toisin sanoen tapahtuma M sisältyy A:han, eli $M \cap A = M$. Umpimähkäisessä eli symmetrisessä valinnassa on $P(M \mid A) = 1/100$. Kertolaskukaavan (Tuomisen lause 1.7.4) mukaisesti

$$P(M) = P(M \cap A) = P(A)P(M \mid A) = \frac{1}{4} \cdot \frac{1}{100} = \frac{1}{400}.$$

(b) Merkitään Villen valintaa tapahtumalla V. Lasketaan kuten a-kohdassa

$$P(V) = P(D)P(V \mid D) = \frac{1}{4} \cdot \frac{1}{500} = \frac{1}{2000}.$$

Vaikka sekä alue että alueen asukas valittiin umpimähkään (symmetrisesti), ei lopputulos ole symmetrinen asukkaiden kesken.

7. Miten edellisessä tehtävässä olisi alueen valinta pitänyt tehdä, jotta jokaisella asukkaalla olisi yhtä suuri todennäköisyys tulla valituksi?

Ratkaisuehdotus:

Siten, että alueen tn tulla valituksi on suhteessa asukaslukuun, eli pitää olla $P(A) = \frac{1}{10}$, $P(B) = \frac{1}{10}$, $P(C) = \frac{3}{10}$ ja $P(D) = \frac{5}{10}$. Tällöin jokaisella asukkaalla on 1/1000 todennäköisyys tulla valituksi.

8. Eräässä populaatiossa on $n = 1000$ henkilöä. Heistä 300 on rokotettu erästä sairautta vastaan. Rokotetuista 3 sairastui ja rokottamattomista 100 sairastui. Tarkastelemme nyt populaatiosta umpimähkään valittua henkilöä X (ts. valintamme osuu kuhunkin henkilöön todennäköisyydellä $1/n$).

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

(b) Mikä on tn, että X sairastui **ja** hänet on rokotettu?

(c) Mikä on tn, että X on rokotettu, ehdolla, että hän sairastui?

Ratkaisuehdotus:

Merkitään R="rokotettu" ja S="sairastui". Koska henkilö valitaan umpimähkään, kyseessä on symmetrinen todennäköisyys, ja tn henkilön kuulua johonkin joukkoon saadaan laskemalla, montako alkiota joukossa on.

(a) Joukossa $S \cap R$ on 3 alkiota (henkilöä), joten $P(S \cap R) = 3/1000$. Toisaalta $P(R) = 300/1000$. Ehdollisen tn:n määritelmästä saadaan

$$P(S \mid R) = \frac{P(S \cap R)}{P(R)} = \frac{3/1000}{300/1000} = \frac{3}{300} = \frac{1}{100}.$$

(b) A-kohdassa jo laskettiin $P(S \cap R) = 3/1000$.

(c) Sairastuneita on $3+100 = 103$, joten sairastumisen todennäköisyys on $P(S) = 103/1000$. Ehdollinen tn lasketaan kuten a-kohdassa, mutta jakaja on eri luku:

$$P(R \mid S) = \frac{P(R \cap S)}{P(S)} = \frac{3/1000}{103/1000} = \frac{3}{103}.$$