Johdatus tilastolliseen päättelyyn
Harjoitus 5 (26. 4.–30. 4. 2010)

1. Tarkastelemme perusjoukkoa Ω, sen osajoukkoja $A, B, A_1, A_2, \ldots$ jotka ovat tapahtumia sekä todennäköisyyttä P. Oletamme, että $P(B) > 0$. Todista, että (tapahtumien joukossa määritelty) ehdollinen todennäköisyys $A \mapsto P(A \mid B) = P(A \cap B)/P(B)$ on todennäköisyys tarkistamalla seuraavat ominaisuudet

   a) $P(A \mid B) \geq 0$ kaikille tapahtumille A.

   b) $P(\Omega \mid B) = 1$.

   c) Jos A_1 ja A_2 ovat erillisiä tapahtumia (ts. $A_1 \cap A_2 = \emptyset$), niin

   $$P(A_1 \cup A_2 \mid B) = P(A_1 \mid B) + P(A_2 \mid B).$$

   Huomautuksia: ehtoa c) kutsutaan nimellä äärellinen additiivisuus; myös täysadditiivisuus (Tuomisen aksiooma TN_3) pätee ehdolliselle todennäköisyydelle. Näiden tarkistusten jälkeen voimme käyttää kaikkia ei-ehdolliselle todennäköisyydelle johdettuja laskukaavoja myös ehdolliselle todennäköisyydelle!

2. Olkoot A, B ja C tapahtumia siten, että $P(A \cap B \cap C) > 0$. Tarkista, että

   a) kertolaskukaava yleistyy kolmelle tapahtumalle muodossa

   $$P(A \cap B \cap C) = P(A)\, P(B \mid A)\, P(C \mid A \cap B),$$

   b) ehdollisen todennäköisyyden kertolaskukaava (kahdelle tapahtumalle) on voimassa, eli

   $$P(A \cap B \mid C) = P(A \mid C)\, P(B \mid A \cap C).$$

   (Muistathan, että todennäköisyys on monotoninen: inkluusiosta $A_1 \subset A_2$ seuraa aina $P(A_1) \leq P(A_2)$. Tämän takia tehtävässä tarvittavat ehdolliset todennäköisyydet ovat hyvin määriteltyjä, sillä $0 < P(A \cap B \cap C)$ ja leikkaus $A \cap B \cap C$ sisältyy kaikkiin tehtävän ehtotapahtumiin.)

   Huom. Tehtävissä 3 ja 4 satunnaismuuttujilla X, Y ja K on diskreetti jakauma, ja niiden yhteis-ptnf toteuttaa luentojen s. 39 **positiivisuusoletuksen**:

   $$P(X = x, Y = y, K = k) > 0 \quad \text{kaikilla } x, y, k.$$

   Sana "kaikilla" tarkoittaa tässä (ja tehtävässä 4) kaikilla $x \in S_X$, kaikilla $y \in S_Y$ ja kaikilla $k \in S_K$. Tässä S_X on satunnaismuuttujan X arvojoukko (ts. X voi saada arvoja vain joukosta S_X), S_Y on Y:n arvojoukko ja S_K on K:n arvojoukko. S_X, S_Y ja S_K ovat kaikki äärellisiä joukkoja.

3. Perustele, miksi funktio

   $$x \mapsto P(X = x \mid Y = y, K = k)$$

   (kiinteillä $y \in S_Y, k \in S_K$) on pistetodennäköisyysfunktio, eli miksi

   $$P(X = x \mid Y = y, K = k) \geq 0 \quad \text{kaikilla } x \in S_X \text{ ja}$$
   $$\sum_{x \in S_X} P(X = x \mid Y = y, K = k) = 1.$$

   (Samanlainen perustelu toimii tietenkin myös muille ehdollisille pistetodennäköisyysfunktioille.)

4. Osoita (positiivisuusoletuksen vallitessa), että X ja Y ovat ehdollisesti riippumattomia ehdolla K jos ja vain jos

$$P(X = x \mid Y = y, K = k) = P(X = x \mid K = k) \quad \text{kaikilla } x, y, k.$$

5. Kulhossa on aluksi 4 palloa, joista K on valkoista ja loput mustia. Ennakokäsityksesi mukaan K:lla voi yhtä todennäköisesti olla mikä tahansa arvoista 0, 1, 2, 3, 4. Nostat silmät sidottuna kulhosta yhden pallon kerrallaan siten, että ennen kutakin nostoa kulhoa ravistetaan perusteellisesti, mutta mitään nostetuista palloista ei palauteta kulhoon. Kunkin noston jälkeen saat tietää nostamasi pallon värin. Kolmessa ensimmäisessä nostossa saat tulokseksi värit valkoinen, musta ja valkoinen. Laske K:n posteriorijakauma.

   Opastus: Koska nostettuja palloja ei palauteta kulhoon, eri nostojen tulokset eivät ole ehdollisesti riippumattomia ehdolla K. Laske uskottavuusfunktio kertolaskukaavan avulla,

$$\begin{aligned} &P(X_1 = 1, X_2 = 0, X_2 = 1 \mid K = k) = \\ &\qquad P(X_1 = 1 \mid K = k)\, P(X_2 = 0 \mid X_1 = 1, K = k)\, P(X_3 = 1 \mid X_1 = 1, X_2 = 0, K = k). \end{aligned}$$

   Tässä esiintyvät ehdolliset todennäköisyydet on helppo järkeillä. (Miten monta palloa ja miten monta valkoista ja miten monta mustaa palloa kulhossa on jäljellä, kun kyseessä oleva ehto on voimassa?)