13.3. Binomijakauma

Kun puhuttiin binomitodennäköisyydestä, sille oli tunnusomaista saman ilmiön toistuminen n kertaa, ja yksittäisen kokeen lopputulosten lukumäärä oli rajattu kahdeksi: sattuu A tai sattuu A:n komplementti $\overline{A}$.

Tapahtuman A esiintymiskertojen lukumäärää (k) voidaan pitää satunnaismuuttujana, joka voi saada arvot 0,1,2,3, ... , n – 1 , n. Kunkin esiintymiskertojen määrän todennäköisyys saadaan lasketuksi lauseen 6.7 nojalla

$$P(\underline{x} = k) = \binom{n}{k} \cdot p^k \cdot q^{n-k}$$

Tapahtuman A esiintymiskertojen lukumäärän sanotaan olevan **binomijakautunut parametrein n ja p.** Tätä merkitään kirjallisuudessa joskus $\underline{x} \sim \mathrm{Bin}(n,p)$.

Binomijakauman odotusarvo tarkoittaa tässäkin jakauman graafisen esityksen massakeskipistettä; mikä on se tapahtuman A esiintymiskertojen lukumäärä, jonka kohdalta tuettuna jakauma pysyisi tasapainossa.

****

LAUSE 8 Jos satunnaismuuttuja $\underline{x}$ on binomijakaantunut parametrein n ja p, niin sen odotusarvo $E\underline{x} = np$ ja keskihajonta $D\underline{x} = \sqrt{npq}$

Tod.: Ei ole virallista kurssia

$$E\underline{x} = \sum_{k=0}^{n} p_k \cdot k = \binom{n}{0} \cdot p^0 \cdot q^n \cdot 0 + \binom{n}{1} \cdot p \cdot q^{n-1} \cdot 1 + \binom{n}{2} \cdot p^2 \cdot q^{n-2} \cdot 2 + \ldots =$$

$$= \sum_{k=1}^{n} \frac{n!}{(n-k)! \cdot k!} \cdot p^k \cdot q^{n-k} \cdot k = np \sum_{k=1}^{n} \frac{(n-1)!}{(n-k)! \cdot (k-1)!} p^{k-1} q^{n-k} =$$

$$= np(q+p)^{n-1} = np, \text{ sillä } p + q = 1.$$

Lukiokurssiin kuulumaton, tai ehkä myöhemmin vastaan tuleva Newtonin binomikaava on ollut yllä kovassa sovellutuskäytössä:

$$(a+b)^n = \binom{n}{0} \cdot a^n b^{n-n} + \binom{n}{1} \cdot a^{n-1} \cdot b + \binom{n}{2} \cdot a^{n-2} b^2 + ... = \sum_{k=0}^{n} \binom{n}{k} \cdot a^k \cdot b^{n-k}$$

Kaavaan on sijoitettu n:n paikalle $n - 1$ ja $a = q$ sekä $b = p$.

Keskihajonnan johto on huomattavasti vaikeampi prosessi, joten se sivuutetaan, mutta löytyy kiinnostuneille alan oppikirjoista.

****

Esim. 1 Eräällä paikkakunnalla suoritettu mielipidemittaus osoittaa, että presidenttiehdokas Hämäläistä kannattaa 70% äänioikeutetuista. Erään kerran mittaustapahtuman aikoihin oli myymäläautoa odottamassa neljä paikkakunnan asukasta. Olkoot $\underline{x}$ satunnaismuuttuja, joka ilmoittaaa mainituista henkilöistä Hämäläistä kannattavien lukumäärän. Laadi $\underline{x}$:n todennäkäisyysjakauma, laske $\underline{x}$:n odotusarvo $E\underline{x}$ ja keskihajonta $D\underline{x}$.

Odotusarvo ja keskihajonta saadaan esitellyn ja osin todistetun lauseen perusteella helposti. Kun ajatellaan yhtä henkilöä kerrallaan, niin tapahtumalle A = {henkilö kannattaa Hämäläistä} on $P(A) = p = 0.7$ ja $p(\overline{A}) = q = 0.3$.

$\underline{x}$	p_i
0	$0.3^4 = \frac{81}{10000} = 0.0081$
1	$4 \cdot 0.7 \cdot 0.3^3 = \frac{189}{2500} = 0.0756$
2	$\binom{4}{2} \cdot 0.7^2 \cdot 0.3^2 = \frac{1323}{5000} = 0.2646$
3	$4 \cdot 0.7^3 \cdot 0.3 = \frac{1029}{2500} = 0.4116$
4	$0.7^4 = \frac{2401}{10000} = 0.2401$

$E\underline{x} = np = 4 \cdot 0.7 = 2.8$

$D\underline{x} = \sqrt{npq} = \sqrt{4 \cdot 0.7 \cdot 0.3} = 0{,}9165151... \approx 0.92$

Jos jakauman satunnaismuuttujan jokaiseen arvoon liittyviä todennäköisyyksiä ei tarvitse laskea ja vaikka tarvitsisikin, odotusarvon ja keskihajonnan laskeminen on huomattavasti helpompaa esitellyn lauseen nojalla kuin suoraan määritelmän mukaan.