Binomikertoimen kaksi tulkintaa

Olkoon joukossa A erilaisia alkioita n kappaletta: $A = \{a_1, \ldots, a_n\}$. Kuinka monta k:n kokoista ($0 \le k \le n$) erilaista osajoukkoa voidaan A:sta poimia? Vastaus on

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{1}$$

erilaista osajoukkoa. Suure $\binom{n}{k}$ on *binomikerroin*. Se luetaan "n yli k:n".

Tulos päätellään näin: Merkitään tuntematonta osajoukkojen lukumäärää N:llä. Kustakin osajoukosta voidaan muodostaa $k!$ järjestettyä jonoa. Tuloperiaatteen mukaan kaikista osajoukoista saadaan $N \times k!$ järjestettyä jonoa. Toisaalta n alkiosta voidaan muodostaa k:n pituisia järjestettyjä jonoja $n!/(n-k)!$. Täytyy päteä $N \times k! = n!/(n-k)!$. Ratkaisemalla yhtälöstä N saadaan tulos (1).

Perustellaan, että binomikerroin (1) on myös on n:n pituisten erilaisten jonojen lukumäärä kahdenlaisista (vaikkapa oransseista ja vihreistä) alkioista o ja v, joita on k ja $n-k$ kappaletta: Merkitään erilaisten jonojen lukumäärää N:llä. Mikäli voitaisiin erotella o-alkiot toisistaan, olisi erilaisia jonoja $N \times k!$ kappaletta, sillä o-alkiot voidaan järjestää $k!$ eri tavalla yhdessä jonossa. Mikäli voitaisiin erotella myös v-alkiot, olisi erilaisia jonoja $N \times k! \times (n-k)!$ kappaletta, sillä v-alkiot voidaan järjestää $(n-k)!$ eri tavalla yhdessä jonossa. Tällöin pystyttäisiin erottelemaan kaikki alkiot, jolloin erilaisia jonoja on $n!$. Näin ollen täytyy päteä

$$N \times k! \times (n-k)! = n!$$

eli

$$N = \frac{n!}{k! \times (n-k)!} = \binom{n}{k}. \tag{2}$$

Binomijakauman perustelu

On tehty tai havaittu n riippumatonta samanlaista Bernoulli-koetta (kussakin tapahtumatodennäköisyys on π). Tällaista yhdistettyä koetta kutsutaan *binomikokeeksi* (*binomial experiment*). Merkitään tapahtumien lukumäärää binomikokeessa y:llä ($0 \le y \le n$). Sitä vastaavan satunnaismuuttujan pistetodennäköisyys on

$$\mathsf{P}(Y = y) = \binom{n}{y} \pi^y (1-\pi)^{n-y}. \tag{3}$$

Jakaumaa kutsutaan *binomijakaumaksi* (otoskoolla n ja parametrilla π) ja sitä merkitään $\mathsf{Bin}(n, \pi)$.

Selitys pistetodennäköisyydelle (3): Kaikkien n:n pituisten tapahtumajonojen, joissa on y tapahtumaa, todennäköisyys on $\pi^y(1-\pi)^{n-y}$. Esimerkiksi jos kolme ensimmäistä koetta onnistuvat, seuraava epäonnistuu ja kaksi viimeistä koetta onnistuu ja epäonnistuu ja onnistumisia on yhteensä y, havaitun tapahtumajonon todennäköisyys on

$$\pi\pi\pi(1-\pi) \times \cdots \times \pi(1-\pi) = \pi^y(1-\pi)^{n-y}.$$

Muoto oikealla saadaan kokoamalla tulon termit. Järjestyksestä riippumatta muoto oikealla pätee, jos onnistumisia on y kappaletta. Vaihtoehtoisia järjestyksiä y:lle onnistumiselle ja $n-y$:lle epäonnistumiselle on binomikertoimen $\binom{n}{y}$

mukainen määrä. Kukin järjestys on erillinen. Todennäköisyys $\mathsf{P}(Y = y)$ saadaan erillisyyden perusteella laskemalla kaikkien mahdollisten jonojen, joissa on y onnistumista, todennäköisyydet yhteen:

$$\mathsf{P}(Y = y) = \pi^y (1-\pi)^{n-y} + \cdots + \pi^y (1-\pi)^{n-y} = \binom{n}{y} \pi^y (1-\pi)^{n-y}.$$