Sotilaallisesta tilastolaskennasta

Yleisesikuntaeversti L. Kaje

JOHDANTO

Puolustuslaitoksen "paperisota" on suurelta osaltaan tilastojen laatimista käytetystä, jäljellä olevasta ja tarvittavasta materiaalista, monenlaisista vahvuuksista, menetyksistä, sairauksista, mielialasta, palvelussuhteista ja monesta muusta, kuten on yleisesti tunnettua. Mitään korkeampaa tilastomatematiikkaa ei tällaisessa työssä yleensä tarvita, vaikka selvyyteen ja yksinkertaistukseen pyrkivä, jatkuva rationalisointi onkin tarpeen.

Huomattava osa etenkin teknillisissä tehtävissä toimivia upseereja joutuu kuitenkin joko säännöllisesti tai tilapäisesti ratkomaan probleemeja, jotka edellyttävät edellistä laajempaa tilastotieteen hyväksi käyttöä. Milloin tarve on tilapäinen, suhteellisen harvoin uusiintuva, ovat vaikeudet usein odottamattoman suuret. Aikaisemmin hankitut pohjatiedot ovat osaksi unohtuneet tai ne puuttuvat miltei kokonaan. Alan kirjallisuudesta saatu tietous on hajanaista ja soveltuu ehkä huonosti sotilaallisiin tehtäviin. Seurauksena saattaa olla tehtävän osittainen laiminlyönti tai kiertäminen, jolloin saavutettu työn tulos lopultakin vähäpätöisestä syystä muodostuu epäluotettavaksi, jopa harhauttavaksikin. Erikoisesti näyttää olevan tarjona vaara käyttää väärin tilastollisissa laskuissa esiintyviä vakioita ja kertoimia.

Tämän kirjoituksen tarkoituksena oli alun perin auttaa ennen kaikkea oikeiden kertoimien löytämisessä, mutta lähemmin tarkasteltuna näytti tarpeelliselta samalla koota suppeassa muodossa yhteen alkeistietoa yleensä todennäköisyyslaskennan ja tilastotieteen perusteista. Tällöin on tosin vaikeata päätellä, mitä pitäisi esittää ja mitä jättää pois, sillä aihepiiri on itse asiassa tavattoman laaja ja nopean kehityksen alainen. Aina on myös tarjona vaara, että pintapuolisia tietoja käytetään väärin ja harkitsematta. Mutta ehkä haitta on vielä suurempi silloin, kun tietojen puuttuessa yksinkertaisetkin keinot jäävät käyttämättä.

Erilaisten testien kehittäminen ja niiden soveltaminen on nykyisin tunkemassa läpi kaikilla aloilla. Näitä ei alempana kuitenkaan käsitellä, sillä se veisi tässä jo liian pitkälle.

Monimutkaiselta näyttäviä merkintätapoja on koetettu välttää, koska ne— ammattimatemaatikosta ehkä käsittämättömällä tavalla — tympäisevät lukijaa, joka ei ole niihin tottunut ja saattavat hänet luulemaan asioita vaikeammiksi kuin ne itse asiassa ovat.

Kirjoituksessa ei millään tavoin puututa siihen, miten tilastollisten menetelmien käyttöä edellyttävä tutkimustyö suoritetaan. Opastusta työn suunnittelussa ja suorituksessa on saatavissa mm Helsingin Yliopiston julkaisusta ”Johdatusta tilastolliseen tutkimukseen”. Siinä on esitetty myös yleisiä ohjeita mm taulukointitekniikasta.

I TILASTOLASKENNALLISIA MÄÄRITELMIÄ JA SUHDELUKUJA

a. Keskiluvut

Tilastoaineiston ominaisuuksia kuvaavien ns tunnuslukujen tärkeän alaryhmän muodostavat keskiluvut. Keskiluku tarkoittaa tavallisimmin aritmeettista keskiarvoa, mutta myös geometrista tai harmoonista keskiarvoa, keskeisarvoa (mediaania) ym riippuen siitä, mikä näistä parhaiten vastaa haetun tilastollisen suureen oikeana pidettävää arvoa. Tällöin on ratkaiseva merkitys lähinnä suureen jakautuman luonteella.

Aritmeettinen (A), geometrinen (G) ja harmooninen (H) keskiarvo positiivisille luvuille $a_1, \ldots, a_n$ lasketaan seuraavasti:

$$A = \frac{a_1 + \ldots + a_n}{n}$$

$$G = \sqrt[n]{a_1 \,.\, a_2 \,.\, \ldots \,.\, a_n}$$

$$\frac{1}{H} = \frac{1}{n} \left(\frac{1}{a_1} + \ldots + \frac{1}{a_n}\right)$$

Aritmeettisen keskiarvon käyttö on yleisintä, ehkä jopa liiankin yleistä. Kun tilastosuureen jakautuma on normaali tai yleensä symmetrinen, keskiarvon täytyy katsoa edustavan suureen oikeana pidettävää arvoa sitä tarkemmin, mitä runsaamman havaintomäärän perusteella se lasketaan. Samalla keskiarvo yhtyy keskeisarvoon sekä moodiin eli tyyppiarvoon (empiirisestä jakautumasta puheen ollen käytetään myös nimitystä valta-arvo), ts arvoon, jonka todennäköisyys on suurin. Jakautuman (koko perusjoukkoa edustavaa) keskiarvoa sanotaan myös odotusarvoksi. (Painotetusta keskiarvosta ks "Johdatusta tilastolliseen tutkimukseen".) Suureen todellinen arvo saattaa olla ennakolta tunnettu. Tarkassa koordinaatistossa toimittaessa on esim ampumaetäisyys käytännöllisellä tarkkuudella tiedossa ennen ammuntaa, mutta valmisteltu matka on epävarma. Tarkistusammunnan ja korjausten jälkeen tai samassa yhteydessä on keskiarvosääntöä käyttäen halutulla tarkkuudella todettavissa, onko todella saavutettu myös oikea matka.

Geometrinen keskiarvo merkitsee syvemmin käsitettynä oikeastaan sitä, että otetaan tavallinen aritmeettinen keskiarvo, ei itse luvuista, vaan niiden logaritmeista. Geometrinen keskiarvo antaa useissa tehtävissä paremman tuloksen kuin aritmeettinen keskiarvo. Otamme esimerkin.

Järjestelyosaston päällikkö antaa useammalle toimistoupseerille tehtäväksi toisistaan riippumatta suunnitella ryhmän kokoonpanon tiettyä tehtävää varten. Olettakaamme, että ennakolta tunnettu optimivahvuus on 10. Ehdotukset hajaantuvat, mikäli tehtävä on oikein orientoitu, ilmeisesti tämän arvon ympärille ja optimiarvoa suurempia lukuja esitetään keskimäärin yhtä paljon kuin tätä pienempiäkin. Mutta poikkeamat alaspäin ovat varmuudella pienempiä kuin poikkeamat ylöspäin, koska arvo nolla on rajoittavana tekijänä. On yhtä todennköistä, että joku esittää vahvuudeksi lukua 5 (puolet optimista) kuin että joku toinen esittää sille lukua 20 (optimi siitä puolet). Näiden geometrinen keskiarvo, mikä tässä yleensäkin sopii käytettäväksi, antaa $\sqrt{5 \cdot 20} = 10$, siis haetun optimiarvon. Aritmeettinen keskiarvo taas on 12.5. Keskiarvojen erosta (tässä 2.5) seuraa, että organisaatiot pyrkivät keskimäärin paisumaan liian suuriksi.

Edellä olemme valinneet käyttöön geometrisen keskiarvon "havaintoarvojen" jakautuman luonteesta johtuen. E h k ä t a v a l l i s i n t a o n g e o m e t r i s e n k e s k i a r v o n k ä y t t ö k u i t e n k i n s i l l o i n, k u n o n k y s e s u h d e l u v u i s t a. Ajatus selvenee parhaiten esimerkin valossa. Olkoon tutkimusperusteilla (1) tultu siihen tulokseen, että organisaatio A on k_1 kertaa niin tehokas kuin organisaatio B. Tutkimusperusteilla (2) on päädytty vastaavaan kertoimeen k_2. Jos k_1 ja k_2 ovat samanveroisia, on meidän — ellemme muuta tiedä — otettava lopulliseksi kertoimeksi $k = \sqrt{k_1 \cdot k_2}$. Tämä johtuu siitä, että suhdeluvun tunteminen ei sano vielä mitään ao suureiden absoluuttisista arvoista, mutta näiden arvojen logaritmien erotus on kylläkin selvillä. Jos esim

$$\frac{a_1}{b_1} = k_1 \text{ ja } \frac{a_2}{b_2} = k_2, \text{ on } \log a_1 - \log b_1 = \log k_1 \text{ sekä}$$

$\log a_2 - \log b_2 = \log k_2$. Ainoa mielekäs keskiarvo tämän perusteella yleensä on $\log k = \dfrac{\log k_1 + \log k_2}{2}$ ja siitä seuraa $k = \sqrt{k_1 \cdot k_2}$.

Jos suhdeluvut tai muut vertausluvut (esim vahvuudet organisaatioesimerkissämme) eroavat vain vähän toisistaan, on käytännössä yhdentekevää käytetäänkö aritmeettista tai geometrista keskiarvoa.

H a r m o o n i s e n k e s k i a r v o n käytöstä otamme seuraavan esimerkin. Kolonna liikkuu puolet matkasta päätietä pitkin nopeusmittarin näyttäessä 60 km/t ja toisen puolen sivutiellä nopeuden ollessa 30 km/t. Keskinopeus v_k saadaan seuraavasti:

$$\frac{1}{v_k} = \frac{1}{2}\left(\frac{1}{60} + \frac{1}{30}\right)$$

Tästä saamme $v_k = 40$ km/t, kun taas tavallinen keskiarvo on 45 km/t, mikä on virheellinen. Pienet nopeudet vaikuttavat lopputulokseen enemmän kuin suuret. Esim pilotpallon nousunopeuskeskiarvoja tilastojen perusteella laskettaessa on välttämätöntä käyttää harmoonisia keskiarvoja. Nopeuksista lasketun tavallisen keskiarvon mielekkyyttä on yleensäkin syytä aina epäillä.

Harmoonisen keskiarvon sijasta voidaan em tehtävissä käyttää tavallista keskiarvoa, mutta laskuperusteena tulee silloin olla ajat (ja matkat), ei nopeudet.

Em kaava voidaan yleispätevämmässä muodossa kirjoittaa painotettuna seuraavasti:

$$\frac{1}{v_k} = k_1 \cdot \frac{1}{v_1} + k_2 \cdot \frac{1}{v_2} + \ldots + k_n \cdot \frac{1}{v_n}$$

missä kertoimen merkitys selvenee parhaiten seuraavasta laajennetusta esimerkistä. Autolla ajettaessa näytti nopeusmittari ensimmäisillä 2 km:llä (x_1) 60 km/t, seuraavilla 5 km:llä (x_2) 75 km/t ja viimeisillä 3 km:llä (x_3) 30 km/t. Kysymme keskinopeutta (v_k) ajetulla 10 km matkalla (x). Kertoimet ovat

$$k_1 = \frac{x_1}{x} = 0.2,\ k_2 = \frac{x_2}{x} = 0.5 \text{ ja } k_3 = \frac{x_3}{x} = 0.3$$

Siis $$\frac{1}{v_k} = \frac{0.2}{60} + \frac{0.5}{75} + \frac{0.3}{30} \text{ ja } v_k = 50 \text{ km/t}$$

K e s k e i s a r v o (mediaani, keskusarvo) on parittomien lukujen keskimmäinen arvo tai parillisten lukujen kahden keskimmäisen luvun keskiarvo. Edellä esitetyssä, geometrisen keskiarvon käyttöä valaisevassa esimerkissämme (ryhmän vahvuuden laskeminen) voisimme käyttää valintaperusteena myös keskeisarvoa. Milloin keskeisarvo ja moodi yhtyvät, mutta keskiarvo sattuu eri kohtaan, voi keskeisarvon käyttäminen (kuten esimerkissäkin) olla edullista. Keskeisarvoa käytetään mm. iskemäkeskipisteen likiarvoisessa määrittämisessä laskujen nopeuttamiseksi ja tarkoituksella eliminoida reunaiskemien ja erityisesti suoranaisten harhautumien keskiarvoa huonontava vaikutus.

On yritetty keksiä keinoja aritmeettista keskiarvoa paremman keskiluvun määrittämiseksi vedoten mm. siihen, että keskiarvoa laskettaessa reuna-arvot saavat saman painon kuin keskellä olevat arvot, vaikka ensiksi mainittujen todennäköisyys on pienempi mutta niiden keskiarvoa horjuttava vaikutus sitävastoin hyvin suuri. Tällöin joudutaan asettamaan normaalijakautuma sinänsä kyseenalaiseksi, sillä sitä johdettaessa lähdetään jo olettamuksesta keskiarvon paremmuuteen. Keskeisarvon käyttö merkitsee tavallaan sellaisen painotetun keskiarvon laskemista, jolloin keskimmäinen arvo saa painon yksi (tai kaksi keskimmäistä kumpikin painon 0.5) ja muut painon nolla. Arvojen todennäköisyyksiin perustuva painotus (reuna-arvoille asteittain pienenevä paino) merkitsee jo vähemmän radikaalista muutosta.

Keskilukua laskettaessa on syytä mahdollisuuksien mukaan ensin tarkastella jakautuman laatua luottamatta liian sokeasti sen normaalisuuteen. Iskemäkeskipistettä nopeallakin tavalla määritettäessä saattaisi tulla kyseeseen muutaman keskimmäisen iskemän keskiarvon käyttö keskeisarvon sijasta. Emme kuitenkaan voi tässä tarkemmin puuttua tämän sotilaallisesti tärkeän kysymyksen pohdintaan.

b. Suhdelukuja

Jos normaalijakautuman vallitessa aina (yksityistapauksia tarkemmin harkitsematta) käytetään keskiarvoa tai aina keskeisarvoa, osoittautuu edellinen tapa ehdottomasti paremmaksi. Koska keskeisarvon sotilaskäyttö on erittäin yleistä, on tärkeätä tuntea asia tarkemmin kuin pelkkänä mainintana. Esitämme sen vuoksi alla T Hojon laskeman ja asiaa valaisevan taulukon (Biometrika, XXIII v 1931, s 315). Siinä on vasemmalla tapausten (esim iskemien) luku (n) ja oikealla keskeisarvon ja keskiarvon keskivirheiden suhdeluku (k_n).

Parittomille ja parillisille tapausluvuille muodostuu kummallekin omat sarjansa. Parillisten sarja on edullisempi, koska keskiarvosääntö vaikuttaa siinä osaksi mukana.

TAULUKKO 1

Keskeisarvon keskivirheen suhde keskiarvon keskivirheeseen (k_n) tapausluvusta (n) riippuvana.

n		k_n	n		k_n
1	1.000		17	1.238	
2		1.000	18		1.210
3	1.160		19	1.239	
4		1.092	20		1.214
5	1.198		25	1.243	
6		1.135	30		1.227
7	1.214		35	1.246	
8		1.160	40		1.233
9	1.223		45	1.247	
10		1.177	50		1.237
11	1.229		55	1.249	
12		1.190	60		1.240
13	1.233		70		1.242
14		1.198	80		1.243
15	1.235		90		1.244
16		1.205	100		1.245
			∞		$1.2533 = \left(\frac{\pi}{2}\right)$

Esim. Olkoon yhden iskemän standardipoikkeama ampumasuunnasta $= a$, iskemäkeskipisteen likiarvon, estimaatin eli poikkeamien keskiarvon keskivirhe siis $= a/\sqrt{n}$, kun n on iskemäluku. Jos n on 8, saamme keskeisarvon virheeksi $1.160 \cdot \frac{a}{\sqrt{8}} \approx 0.41$ a. Jos iskemiä on 9 kpl, saamme nytkin vastaavalle virheelle arvon $1.223 \ \frac{a}{3} \approx 0.41$ a. Parittomasta laukauksesta ei siis ole käytännöllistä hyötyä. (Ks tarkemmin kirjoittajan tutkimustyötä ”Tarkistusammunnan teoria”.)

Keskeisarvon nopeaan tarkistukseen pienillä iskemäluvuilla sopii seuraava menetelmä. Määritetään ns hajonta-alueen keskus mittaamalla (silmämääräisestikin) piste, joka on tasan pisimmän ja lyhimmän iskemän puolivälissä. Näin saadun iskemäkeskipisteen likiarvon keskivirheen suhde keskiarvon keskivirheeseen selvenee taulukosta 2.

TAULUKKO 2

Reunaiskemien keskiarvona määritetyn iskemäkeskipisteen keskivirheen suhde keskiarvon keskivirheeseen (k'_n) iskemäluvusta (n) riippuvana. Sulkeissa vertauksen vuoksi vastaavat arvot taulukosta 1.

n	k'_n	(kn)
1	1.000	(1.000)
2	1.000	(1.000)
3	1.042	(1.160)
4	1.092	(1.092)
5	1.142	(1.198)
6	1.190	(1.135)
10	1.362	(1.177)
20	1.691	(1.214)
1000	7.858	(1.253)

Tapausten luvulla 4 on siis yhdentekevää käytetäänkö keskimmäisten vai äärimmäisten iskemien (vast) keskiarvoja. Suuremmilla parillisilla tapausluvuilla on keskeisarvo tarkempi. Parittomilla tapausluvuilla on reunaiskemien keskiarvo keskeisarvoa tarkempi, kun n on 3 ja 5, suunnilleen yhtä tarkka, kun n on 7, mutta sen jälkeen tulee keskeisarvo ehdottomasti paremmaksi.

Esitämme vielä Chauvenet'n kriterion reuna-arvojen poisjättämiselle. Lasketaan ensin keskiarvo siten, että myös epäiltävä reuna-arvo on laskuissa mukana. Sen jälkeen tarkastellaan tämän epäiltävän havaintoarvon poikkeamaa keskiarvosta normaalijakautuman todennäköisyyslukujen perusteella. Jos kyseisen poikkeaman todennäköisyys on pienempi kuin 1/(2 n), missä n on havaintojen luku, hyljätään tuo epäiltävä arvo. Hylkääminen on aina oikeutettu, jos poikkeama on suurempi kuin 4,5 todennäköistä poikkeamaa (virhettä) eli 3,1 standardipoikkeamaa. Yhden standardipoikkeaman alueelle keskiarvon kummallekin puolen mahtuu normaalijakautumassa 68 % kaikista poikkeamista, kahden standardipoikkeaman alueelle vastaavasti 96 % kaikista poikkeamista ja kolmen alueelle n. 100 % (ulkopuolelle jää vähemmän kuin 0.3 %). Em lukujen perusteella on myös tyydyttävästi pääteltävissä, noudattavatko havaitut poikkeamat normaalijakautumaa. Tarkempia arvoja normaalijakautumasta esitetään jäljempänä.

II YLEISTÄ TODENNÄKÖISYYSLASKENNASTA

a. Kombinaatioteoriaa

Voimme muodostaa 3 alkiosta jonoja eri järjestyksessä $1 \cdot 2 \cdot 3 = 6$:lla eri tavalla (abc, acb, bac, bca, cab, cba). Nimitämme näitä vaihteluiksi, permutaatioiksi. Yleisesti merkitsemme $1 \cdot 2 \cdot 3 \cdot \ldots \cdot (n\text{-}1) \cdot n = n!$ (n kertoma).

Otamme esim k_1 kpl punaisia palloja, k_2 kpl sinisiä jne ja k_p kpl mustia palloja, yhteensä n palloa sijoittaen ne samaan rasiaan. Muodostamme aluksi koko pallomäärästä edellisen mukaisesti erilaisia ryhmityksiä, joiden kokonaismäärä siis on n!. Pidämme kuitenkin nyt samana tapauksena kaikkia niitä järjestyksiä, joissa punaiset pallot ovat samoissa asemissa (esim ensimmäisenä, kolmantena, seitsemäntenä jne) ja sama koskee vastaavasti muita värejä. Erilaisten tapausten kokonaismäärä on tällöin

$$\frac{n!}{k_1! \cdot k_2! \cdot \ldots \cdot k_p!} \qquad k_1 + k_2 + \ldots + k_p = n \quad \text{(vrt multinominen jakautuma)}$$

Jos värejä on vain kaksi, toista väriä k kpl ja toista siis (n—k) kpl, saamme lukumääräksi vastaavasti

$$\frac{n!}{k! \cdot (n-k)!} \quad \text{(vrt binomijakautuma)}$$

Viimeksi esitettyyn tulokseen tulemme toistakin tietä. Oletamme, että meillä on n kpl muuten samanlaisia mutta numeroituja palloja, joten ne siis ovat kaikki toisistaan eroitettavissa. Nyt kysymme, montako yhdistelyä, kombinaatiota, on muodostettavissa poimimalla rasiasta kerrallaan k palloa siten, että ainakin yksi pallo joka kerta vaihtuu. Siitä, missä järjestyksessä pallot kussakin ryhmässä ovat, emme välitä. Näiden kombinaatioiden määrä on

$$\frac{n!}{k! \cdot (n-k)!} = \binom{n}{k} \qquad \text{(lue: n k:n yli)}$$

$$= \frac{n\,(n-1) \cdot \ldots \cdot (n-k+1)}{k!}$$

$$= \frac{n\,(n-1) \cdot \ldots \cdot (k+1)}{(n-k)!}$$

Pane merkille, että

$$\binom{n}{k} = \binom{n}{n-k}$$

Esim viiden kombinaatiot kaksittain ovat $5 \cdot 4/2 = 10$ ja kolmittain $5 \cdot 4 \cdot 3/2 \cdot 3 = 10$.

Tulos on edelleen sama, jos merkitsemättömät ja toisensa korvaavat alkiot sijoitetaan paikkoihin, jotka on "numeroitu". Esim 3 palloa voidaan sijoittaa 5:een numeroituun koloon 10 eri tavalla.

Myös ns muuntelut, variaatiot, saattavat tulla kyseeseen. Tällöin otamme n kpl:een alkiojoukosta kerrallaan k alkiota ja laskemme näiden alkioiden kaikki permutaatiot, joita on $n!/(n-k)!$, kun kaikki k alkiota käsittävät erilaiset ryhmät $\left(\frac{n!}{k!\,(n-k)!}\ \text{kpl}\right)$ on otettu huomioon.

Jos n on suuri, n! voidaan likimäärin arvioida Stirlingin epäyhtälön avulla:

$$\sqrt{2\pi} \cdot n^{n+½} \cdot e^{-n} < n! < \sqrt{2\pi} \cdot n^{n+½} \cdot e^{-n} \left(1 + \frac{1}{4n}\right)$$

missä e on luonnollisen logaritmijärjestelmän kantaluku.

b. Todennäköisyyden laskeminen

Todennäköisyyksiä laskettaessa menetellään periaatteessa siten, että suotuisien tapausten luku pannaan todennäköisyyttä ilmaisevan luvun osoittajaksi ja kaikkien mahdollisten tapausten luku nimittäjäksi. Olettakaamme esim, että meillä on kori, johon on sijoitettu 3 punaista palloa ja 4 mustaa palloa. Nostamme korista umpimähkään 5 palloa pois. Nyt kysymme, millä todennäköisyydellä näistä nostetuista palloista on 2 punaista (ja 3 mustaa). Laskemme suotuisat tapaukset seuraavasti. Kolmesta punaisesta pallosta saamme kahden ryhmiä $\binom{3}{2}$ kpl $= 3$ kpl. Neljästä mustasta taas syntyy kolmen ryhmiä $\binom{4}{3}$kpl $= 4$ kpl. Suotuisten tapausten luku on kaikkiaan ilmeisesti $3 \cdot 4 = 12$. Kaikki nostomahdollisuudet taas saamme laskemalla koko pallomäärän eli seitsemän kombinaatiot viisittäin, siis $\binom{7}{5} = \binom{7}{2} = 21$. Haettu todennäköisyys on siis $\frac{12}{21}$. Olemme tässä erikoistapauksessa johtaneet erään ns hypergeometriseen jakautumaan (ks alempana) liittyvän todennäköisyyden.

Erikoistapauksissa on ilmiön todennäköisyyden laskuperiaatetta sovellettava tilanteen mukaisesti. Otamme esimerkin. Jollakin paikkakunnalla sataa keskimäärin 2 tuntia päivässä. Sateen todennäköisyys on ilmeisesti $\frac{2}{24} = \frac{1}{12}$. Tapausten luvut on tässä siis korvattu niitä vastaavilla ajoilla.

Olettakaamme, että meidän on määritettävä ilmiön todennäköisyys koesarjojen perusteella. On suoritettu n koetta ja todettu, että tutkittava ilmiö on esiintynyt näissä kokeissa m kertaa. Hakemamme todennäköisyyden likiarvo on ilmeisesti $\frac{m}{n}$. Nyt voimme 0,5 prosentin erehtymismahdollisuudella sanoa, että haettu todennäköisyys on rajojen

$$\frac{m}{n} \pm \frac{1.5}{\sqrt{n}}$$

välissä (varmuus 0.995). Jos rajat eivät tyydytä, on lisättävä koesarjoja. Tarkkuus paranee (rajat supistuvat) suhteessa havaintoluvun neliöjuureen. Satakertainen havaintomäärä lisää kymmenkertaisesti tarkkuutta. Niin suuri varmuus kuin edellä on esitetty, johtaa käytännössä yleensä suuriin koesarjoihin.

c. Tavallisimpia todennäköisyyskaavoja

Voimme esim merkitä

$$t\{x+y\} = t\{x\} + t\{y\},\ \text{kun}\ t\{x, y\} = 0$$

Tässä tarkoittaa $t\{x\}$ todennäköisyyttä, että ilmiö x esiintyy ja $t\{y\}$, että ilmiö y esiintyy, ja $t\{x, y\}$ todennäköisyyttä, että kumpikin esiintyy yht'aikaa. Koska viimeksi mainittu on merkitty nollaksi, ovat x ja y toisensa poissulkevia, ts jos toinen esiintyy niin toinen ei esiinny. Todennäköisyys $t\{x+y\}$ taas tarkoittaa, että x tai y tai molemmat esiintyvät. Tässä jää viimeinen mahdollisuus pois annetun lisäehdon takia. Jos ilmiöt eivät ole toisiaan poissulkevia, vaan riippumattomia, saadaan

$$t\{x+y\} = t\{x\} + t\{y\} - t\{x, y\}\ \text{ja}$$
$$t\{x, y\} = t\{x\} \cdot t\{y\}$$

Esim 1. Todennäköisyys sille, että nopan heitossa esiintyy ykkönen tai kakkonen (toisensa poissulkevat tapaukset) on $\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$

Esim. 2. Todennäköisyys sille, että maalia ammuttaessa isku menee sivusuunnassa ohi, olkoon 10 % ja että se menee pituussuunnassa ohi, samoin 10 %. Todennäköisyys sille, että isku menee joko sivu- tai pituussuunnassa ohi, on $0{,}1 + 0{,}1 - 0{,}1 \cdot 0{,}1 = 0{,}2 - 0{,}01 = 0{,}19$ (riippumattomat tapaukset).

Sotilaallisissakin tehtävissä voimme joutua käsittelemään todennäköisyyksiä, joihin liittyy erilaisia e h t o j a. Otamme asiaa valaisevan esimerkin. Huoltopäällikkö saattaa sodassa todeta, että joukot joutuvat usein puutteellisesti varustettuina marssimaan sateessa, ja ryhtyy harkitsemaan, olisiko asiassa ryhdyttävä joihinkin toimenpiteisiin. Tätä varten hän tarvitsee erinäisiä lisätietoja.

Todennäköisyysopilliseen muotoon puettuina vastausta kaipaavat kysymykset saattaisivat olla:

a) millä todennäköisyydellä joukot ovat marssilla, kun sataa,

b) millä todennäköisyydellä sataa, kun joukot ovat marssilla,

c) mikä on todennäköisyys sille, että joukot ovat marssilla ja sataa.

Olettakaamme, että huoltopäällikön tiedossa ovat seuraavat seikat:

Kyseisen taistelutoiminnan aikana (mikä on jo ehto sinänsä) joukkojen ajasta kuluu marsseihin $\frac{1}{3}$ ja muuhun toimintaan $^2/_3$. Siis todennäköisyys sille, että joukot ovat marssilla, on $t\{x\} = \frac{1}{3}$. Toisaalta kyseisenä vuodenaikana ja kyseisellä alueella (ehtoja nämäkin, vaikka eivät vaikuta laskukaavoihin) on sadetta $\frac{1}{4}$ vuorokaudesta ja merkitsemme tämän sateen todennäköisyyden $t\{y\} = \frac{1}{4}$. Lisäksi on osoittautunut, että joukot ovat sateisella säällä 2 kertaa niin useasti marssilla kuin poudalla johtuen erinäisistä taktillisista syistä.

Otamme käyttöön seuraavat uudet merkinnät:

Todennäköisyys, että joukot (eivät suinkaan välttämättä kaikki yht'aikaa) ovat marssilla juuri silloin kun sataa, siis sateella suoritettujen marssiaikojen suhde koko sadeaikaan, olkoon $t\{x|y\}$. Todennäköisyys, että juuri silloin sataa, kun joukot ovat marssilla, siis sateella suoritettujen marssiaikojen suhde koko marssiaikaan olkoon taas $t\{y|x\}$. Vihdoin todennäköisyys sille, että marssitaan ja sataa, eli sademarssiaikojen suhde koko toiminta-aikaan olkoon $t\{x, y\}$. Vertailuyksikkönä voi olla tietenkin esim. vuorokausi tai viikko, se ei vaikuta tulokseen. Kun edellä olemme symbolisoineet marssimista kirjaimella x, merkitsemme sen vastakohtaa, siis ei-marssimista $\overline{x}$, siis panemalla viivan päälle. Esim merkintä $t\{y \mid \bar{x}\}$ tarkoittaa todennäköisyyttä, että sataa silloin, kun ei olla marssilla. Ei-sadetta merkitsemme vastaavasti $\bar{y}$. Seuraavat yleiset kaavat mm pätevät:

$$t\{x,y\} = t\{x\} \cdot t\{y \mid x\} = t\{y\} \cdot t\{x \mid y\}$$

$$t\{x\} = t\{y\} \cdot t\{x \mid y\} + t\{\bar{y}\} \cdot t\{x \mid \bar{y}\}$$

Jos, mikä tässä ei pidä paikkaansa, x ja y ovat riippumattomia, saisimme vielä

$$t\{x,y\} = t\{x\} \cdot t\{y\}$$

mikä aikaisemmin on jo esitetty.

Tiedämme, että $t\{x\} = 1/3$ ja $t\{y\} = 1/4$. Ilmeisesti $t\{\bar{y}\} = 1 - t\{y\} = 3/4$ ja edelleen $t\{x \mid y\} = 2\, t\{x \mid \bar{y}\}$, koska sateessa marssitaan 2 kertaa niin usein kuin poudalla. Näin ollen $t\{x\} = 1/3 = 1/4 \cdot 2\, t\{x \mid \bar{y}\} + 3/4\, t\{x \mid \bar{y}\}$ ja siis $t\{x \mid \bar{y}\} = 4/15$ sekä edelleen $t\{x \mid y\} = 8/15$. Näin ollen $t\{x, y\} = t\{y\} \cdot t\{x \mid y\} = 1/4 \cdot 8/15 = 2/15$ ja $t\{y \mid x\} = \dfrac{t\{x, y\}}{t\{x\}} = 2/5$.

Siis: a) Todennäköisyys sille, että marssitaan kun sataa $(t\{x \mid y\})$ on $\dfrac{8}{15}$ ja kun sadetta on 6 t/vrk, se tekee 3,2 t/vrk.

b) Todennäköisyys, että sataa, kun joukot ovat marssilla $(t\{y \mid x\})$ on $\dfrac{2}{5}$, mikä 8 marssitunnista tekee sekin 3,2 t/vrk.

c) Todennäköisyys sille, että joukot ovat marssilla ja sataa $(t\{x,y\})$ on $\dfrac{2}{15}$ mikä 24 tunnista on jälleen 3,2 t/vrk.

Laskut ovat siis ilmeisesti oikeat.

Jos marssitaan tarpeen mukaan, satoi tai paistoi, on $t\{x,y\} = t\{x\} \cdot t\{y\} - \dfrac{1}{3} \cdot \dfrac{1}{4} = \dfrac{1}{12}$, joten sateessa marssimista olisi silloin 2 tuntia päivässä.

III TAVALLISIMMAT JAKAUTUMAT

a. Hypergeometrinen jakautuma

Satunnaissuureisiin liittyvät todennäköisyydet (alla $t\{x\}$) muodostavat sen jakautuman.

Hypergeometrinen jakautuma määrittyy seuraavasti:

$$t\{x\} = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \qquad x = 0,1, \ldots, m$$

missä m = pienempi luvuista M tai n.

Tämä jakautuma ilmaisee millä todennäköisyydellä $t\{x\}$ määrästä N otetussa n kpl:een otoksessa (näytteessä) joku ilmiö (löydös) esiintyy x kertaa, kun se koko määrässä N esiintyy M kertaa. Esim pussissa on 1000 arpaa (N = 1000) ja niistä voittoarpoja 100 kpl (M = 100). Henkilö ostaa 10 arpaa (n = 10). Todennäköisyys sille, että niistä esim 2 kpl (x = 2) on voittoja, on

$$t\{2\} = \frac{\binom{100}{2}\binom{1000-100}{10-2}}{\binom{1000}{10}} \approx 0{,}2$$

Laskeminen on varsin hankala tehtävä. Suurin todennäköisyys sattuu x:n arvolle, joka on lähinnä pienempi kokonaisluku kuin $\frac{(n+1)\ (M+1)}{(N+2)}$

Jakautuman keskiarvo on $E = n\frac{M}{N} = np$, kun merkitsemme $p = \frac{M}{N}$. Niinpä edellä käsitellyssä esimerkissämme saamme: $E = 10 \cdot \frac{100}{1000} = 1$, siis 10 arvan nostolla tulee keskimäärin 1 voitto. Standardipoikkeama on

$$\sigma = \sqrt{n \frac{M}{N} (1-\frac{M}{N}) \quad (1-\frac{n-1}{N-1})}$$

$$= \sqrt{np\,(1-p) \quad (1-\frac{n-1}{N-1})}$$

b. Binomijakautuma

Binomijakautuman todennäköisyydet saadaan em jakautuman rajamuotona, kun annamme suureiden N ja M lähetä äärettömyyttä, kuitenkin niin, että suhde $M : N = p$ pysyy vakiona. Siten saamme

$$t\left\{x\right\} = \binom{n}{x} p^x (1-p)^{n-x}$$

Binomijakautuma ilmaisee todennäköisyyden $t\left\{x\right\}$ sille, että (äärettömän suuresta) perusjoukosta otetussa n kpl:een otoksessa ilmiö esiintyy x kertaa, kun sen suhteellinen osuus perusjoukossa on p. Esim. Heitämme rahaa 5 kertaa ($n = 5$) ja kysymme, millä todennäköisyydellä tuloksessa on 2 klaavaa (ja 3 kruunua). Ilmeisesti tässä on $p = ½$. Saamme

$$t\left\{2\right\} = \binom{5}{2} \left(\frac{1}{2}\right)^2 \left(1-\frac{1}{2}\right)^{5-2} = \frac{5}{16}$$

Jos $x = n$, saadaan $t\left\{n\right\} = p^n$.

Jos siis ilmiön todennäköisyys yhdessä kokeessa on p, on todennäköisyys sille, että se esiintyy n kertaa toistetussa kokeessa joka kerta $= p^n$.

Todennäköisyys on suurin sillä x:n arvolla, joka on lähinnä pienempi kokonaisluku kuin $(n + 1)p$. Jos $(n+1)p$ on kokonaisluku, on joko samalla tai yhtä pienemmällä x:n arvolla suurin todennäköisyys. Tämä todennäköisyys on likimäärin, kun n on suuri

$$t\left\{x\right\}_{maks} \approx \frac{1}{\sqrt{2\pi\ np}\ \sqrt{1-p}}$$

Binomijakautuman keskiarvo on $E = np$ ja standardipoikkeama $\sigma = \sqrt{np\,(1-p)}$.

Binomijakautuman likiarvona voidaan silloin, kun $np\,(1-p) > 9$, käyttää hyvin normaalijakautumaa, jonka keskiarvo on np ja standardipoikkeama $\sqrt{np\,(1-p)}$.

Em arvannostamisesimerkki soveltuu myös binomijakautumaan sillä muutoksella, jolla tässä on vain teoreettinen merkitys, että jokainen nostettu arpa pannaan heti noston jälkeen takaisin pussiin, jotta voittoarpojen ja tyhjien arpojen suhde pysyisi muuttumattomana. Tällöin saadaan

$$t\{2\} = \binom{10}{2}\left(\frac{1}{10}\right)^2\left(\frac{9}{10}\right)^8 \approx 0{,}2$$

c. Multinominen jakautuma

Voimme itse asiassa pitää binomijakautumaa myös multinomisen jakautumisen erikoistapauksena. Olkoon esim pussissa punaisia, sinisiä, mustia jne palloja suhteellisten osuuksien ollessa $p_1\, p_2, \ldots, p_k$. Nostamme pussista palloja yhden kerrallaan ja panemme nostetun pallon heti takaisin. Teemme yhteensä n nostoa ja kysymme todennäköisyyttä sille, että saamme x_1 kpl punaisia, x_2 sinisiä jne palloja, jolloin tietenkin $\Sigma x_i = x_1 + x_2 + \ldots + x_k = n$ ja $\Sigma p_i = p_1 + p_2 + \ldots p_k = 1$. Tämä todennäköisyys on

$$t\{x_1, x_2 \ldots, x_k\} = \frac{n!}{x_1! \cdot x_2! \ldots x_k!} p_1^{x_1} \cdot p_2^{x_2} \ldots \cdot p_k^{x_k}$$

$$\Sigma p_i = 1$$
$$0 \leqq x_i \leqq n$$
$$\Sigma x_i = n$$

Esim pussissa on tasan yhtä paljon kolmen värisiä palloja. Todennäköisyys sille, että 3 nostolla jokainen nostettu pallo on eri värinen, on

$$t\{1, 1, 1\} = 3!\left(\frac{1}{3}\right)^3 = \frac{2}{9}$$

d. Poissonin jakautuma

Poissonin jakautumaan voimme vuorostaan päästä binomijakautuman rajamuotona antamalla suureen n lähetä ääretöntä ja p:n nollaa, kuitenkin siten, että binomijakautuman keskiarvo = np pysyy muuttumattomana = λ. Tällöin saamme

$$t\{x\} = \frac{\lambda^x}{x!} e^{-\lambda} \qquad x = 0, 1, \ldots$$

Poissonin jakautuman keskiarvo on $E = \lambda$ ja standardipoikkeama $\sigma = \sqrt{\lambda}$.

Voimmekin soveltaa Poissonin jakautumaa binomijakautuman likiarvona silloin, kun n on suuri ja p pieni. Jälleen em arpomisesimerkkiin sovellettuna saamme ($np = \lambda = 10 \cdot 0{,}1 = 1$)

$$t\{2\} = \frac{1^2}{2!} e^{-1} \approx 0.2$$

Poissonin kaavalla laskettuna (ja sen taulukoita käyttäen) saadaan lopputulos kätevimmin. Suurin todennäköisyys on sillä x:n arvolla, joka on lähinnä pienempi kuin λ.

Kun $\lambda > 9$, voidaan Poissonin jakautuma hyvin korvata normaalijakautumalla, jonka keskiarvo on λ ja standardipoikkeama $\sqrt{\lambda}$.

Poissonin jakautuma on etenkin sotilastehtävissä siksi tärkeä, ettemme voi tyytyä käsittelemään sitä vain binomijakautuman likiarvona ja korvikkeena. Johdamme sen vuoksi Poissonin vakion λ merkityksen erityisesti ammuntaan sovellettuna. Oletamme, että meillä on aluemaali, jonka pinta-ala on A. Tämän alueen sisällä on yksi tai useampia osamaaleja, kukin pinta-alaltaan a (esim pesäkkeitä). Todennäköisyys sille, että tähän aluemaaliin umpimähkään ammuttu iskemä sattuu tiettyyn osamaaliin, on ilmeisesti $\frac{a}{A} = p$. Olkoon ammuttujen ja aluemaalin sisälle osuneiden iskemien luku n. Silloin on

$$\lambda = np = n \frac{a}{A} = \frac{n}{A} \cdot a$$

Koska n on iskemäluku ja A aluemaalin pinta-ala, on $\frac{n}{A}$ ilmeisesti keskimääräinen iskemäluku aluemaalin pintayksikköä kohti, ts keskimääräinen iskemätiheys $= \varrho$ ja tulo $\varrho \cdot a = \lambda$ ilmaisee siis keskimääräisen iskemäluvun osamaalia kohti, jos osamaalin pinta-ala on a.

Jos nyt lähtisimme siitä, että esim tuli-iskuista aina osuu yhtä suuriin aluemaaleihin sama laukausmäärä, se ei ensinnäkään vastaisi todellisuutta, ja toiseksi teholaskut tulisivat (binomikaavalla) mutkikkaiksi, koska ne olisi laskettava erikseen jokaista tapausta varten, jolloin aluemaalin koko muuttuu, vaikka keskimääräinen iskemätiheys ϱ pysyisikin samana (osoitettu tekijän kirjoituksessa sotilasaikakauslehden 6/50 sivulla 39 ja myös alempana). Jos annamme osamäärässä $\frac{n}{A}$ sekä osoittajan että nimittäjän lähetä äärettömyyttä osamäärän arvon silti pysyessä muuttumattomana, tulemme Poissonin jakautumaan. Em täydennys merkitsee sitä, että oletamme ammuttavan aina hyvin suurelle maalialueelle täysin sattumanvaraisesti, joten osumatodennäköisyys kaikkialla on sama (sikäli tasainen jakautuma). Vastaavasti oletamme laukausluvun kasvavan samassa suhteessa kuin aluemaalin kokokin kasvaa, joten iskemätiheys ei keskimäärin muutu. Tämä laajennus tuo mukanaan senkin mahdollisuuden, etä joku kokonainen aluemaali voi harvinaisissa tapauksissa (joita kyllä sattuu) jäädä kenties vaille yhtäkään iskemää ja toisaalta sille voi ainakin teoreettisesti tulla iskemiä enemmän kuin käytännössä on ammuttukaan, minkä todennäköisyys kylläkin tulee häviävän pieneksi. Laskut tulevat näin sikäli yleispäteviksi, että tulokset riippuvat vain keskimääräisestä iskemätiheydestä (ϱ) ja osamaalien koosta (a). Näin tulkitussa Poissonin jakautumassa katsomme siis suureen λ merkitsevän joko sellaisenaan keskimääräistä iskemälukua osamaalin pinta-alaa (vast) kohti tai muodossa $\lambda = \varrho \cdot a$ katsomme ϱ:n merkitsevän iskemätiheyttä pintayksikköä (käytännössä useimmiten hehtaaria) kohti ja a:n osamaalin pinta-alaa (vastaavasti hehtaareissa). (Poissonin jakaantuman pätevyyttä iskemäjakaantuman tulkitsijana on erityisesti tutkinut T Kallio, Tiede ja Ase n:o 15). Poissonin jakautuma kelpaa ainakin erikoismuodossaan myös sirpalejakautuman tutkimiseen (tekijän kirjoitus sotilasaikakauslehdessä 1/51).

Suure x, mikä binomijakautumassa merkitsee varsinaisesti "löydöksiä" (kuten myös Poissonin jakautumassa silloin, kun sitä käytetään binomijakautuman likiarvona), merkitsee nyt tapauksia, jolloin osamaaliin osuu 0, 1, 2 ... iskemää. Koska tämä merkityksen muutos saattaa herättää ihmetystä, on syytä selostaa sitä hieman tarkemmin.

Arpojen nostamista käsittelevässä esimerkissämme oli voittoja 10 % arpojen koko määrästä. Voimme konstruoida esimerkin yksinkertaisimmin siten, että 1000 arvan sijasta käytämme vain 10 arpaa, joista yksi on voittoarpa. Heti arvan nostamisen jälkeen panemme sen takaisin pussiin. Jos siis esim 10 nostolla tulee kaksi voittoa, kuten esimerkissämme, merkitsee se nyt, että sama arpa on vedetty 2 kertaa eli ammuntaan sovellettuna "samaan maaliin on osunut kaksi kertaa". Voimme nim kuvitella, että meillä on esim hehtaarimaali, joka on jaettu 10 osamaaliin ja yksi näistä on merkitty. Että saamme 10 laukauksella ($n = 10$) tähän maaliin 2 osumaa, vastaa tarkoin arpojen nostamista 10 arpaa sisältävästä pussista. Tässä tapauksessa pätee binomijakautuma ja näytämme, että tulos riippuu aluemaalin koosta, kuten aikaisemmin on väitetty. Kun ammumme 10 iskemää 1 ha:lle, on todennäköisyys sille, että merkittyyn maaliin tulee 2 osumaa, binomikaavan mukaan ($p = \frac{1}{10}$)

$$t\left\{2\right\} = \binom{10}{2}\left(\frac{1}{10}\right)^2\left(\frac{9}{10}\right)^8 = \frac{1}{2}\cdot 0{,}9^9 \approx 0{,}2$$

Jos nyt ammumme 2 ha:n aluemaaliin $2\,n = 20$ iskemää, jolloin iskemätiheys ha kohti ei siis keskimäärin muutu, on jälleen binomikaavan mukaan ($p = \frac{1}{20}$)

$$t\left\{2\right\} = \binom{20}{2}\left(\frac{1}{20}\right)^2\left(\frac{19}{20}\right)^{18} = \frac{1}{2}\cdot\left(\frac{19}{20}\right)^{19} = \frac{1}{2}\cdot 0.95^{19} \approx 0{,}2$$

Loppulausekkeiden muodosta näemme, että tuloksilla on eroa, vaikka lopputulos likiarvona onkin sama. Suurentamalla maalialuetta joudumme rajatapauksena Poissonin jakautumaan. Kahden viimeksi esitetyn laskumuunnoksen ero johtuu siitä, että jälkimmäisessä tapauksessa tulevat mukaan nekin osumatodennäköisyydet, jolloin maaliin osuu enemmän kuin n iskemää aina 2 n iskemään asti, kun taas edellisessä tapauksessa suurin mahdollinen määrä oli n iskemää (n = 10). Näin ollen voidaan myös väittää, ettei Poissonin kaavakaan ammuntaan sovellettuna ole suinkaan oikea, sillä jos esim iskussa ei ole enempää kuin 72 laukausta, ei mihinkään osamaaliin voi ainakaan osua sitä enempää. Kuitenkin virhe on häviävän pieni (väärät todennäköisyydet numeroarvoltaan mitättömiä). Edellä on joka tapauksessa samalla myös osoitettu, että jo binomijakautuma on tulkittavissa myös siten, että x siinä ilmaisee esim osumia samaan maaliin, joten siirtyminen siitä Poissonin jakautumaan käy tässä mielessä ymmärrettäväksi.

Binomijakautuma edellyttää rajatonta perusjoukkoa — tahi ylipäänsä sitä, että otos ei vaikuta muutoksia todennäköisyyssuhteisiin. Äskeisessä 10 kpl:een esimerkissämme tämä saavutettiin sillä, että eroitettu arpa pantiin heti takaisin pussiin, tai että tarkasteltu osamaali jäi edelleen alttiiksi tulelle. Mutta kun muutamme tarkastelutapaa ajatellen nimenomaan toistuvaa osumista samaan maaliin, silloin tuo ääretöntä vastaava perusjoukko supistuu 10 kpl:seen osamaaleja ja sama binomikaava pätee vain tässä ainoassa tapauksessa. Siksi siis suoritetaan em "rajankäynti" ja päädytään Poissonin jakautumaan yleistyksen saavuttamiseksi.

Poissonin jakautuma kelpaa binomijakautumassa normaalisti haettujen todennäköisyyksien likiarvoiseen laskentaan (niin ja niin monta voittoa, niin ja niin moneen maaliin osuma, mutta vain yksi) silloin, kun p on pieni, ts että on hyvin epätodennäköistä, että laajalle alalle harvakseltaan suoritettu ammunta antaisi samaan maaliin useampaa kuin yhden osuman. Vain silloin binomijakautuman keskiarvo np vastaa Poissonin jakautuman suuretta λ, ja voidaan käyttää siis λ-taulukoita.

On syytä panna merkille, että silloin kun Poissonin jakautumaa käytetään muussa mielessä kuin binomijakautuman likiarvojen (osuma niin ja niin moneen maaliin) laskemiseen, siis esim ammunnan teholaskemissa (niin ja niin monta osumaa samaan maaliin, vähintään yksi osuma samaan maaliin, niin ja niin moneen prosenttiin maaleista niin ja niin monta osumaa), ei sen soveltamisella ole käytännöllisiä rajoituksia. Iskemätiheyttä siis voidaan nostaa mielin määrin, tutkia patteriston ja suurienkin tykistöryhmien tehoja tulosten siitä huonontumatta. Niin suurien x:n arvojen käyttö (hyvin monta osumaa samaan osamaaliin), että tulokset alkaisivat olla mielettömiä, ei teholaskelmissa tule koskaan kysymykseen.

Koska kaikki osamaalit ovat samassa asemassa, seuraa tästä, että jos esim kahden osuman todennäköisyys yhteen maaliin on 0.1, niin myös 10 % kaikista maaleista ilmeisesti (jos maaleja on paljon) saa nuo 2 osumaa.

Poissonin jakautuman $t\{o\} = e^{-\lambda}$. Tämä on siis todennäköisyys sille, että osamaaliin ei tule yhtään osumaa tai että vastaava murto-osa kaikista maaleista jää vaille osumia. Tästä seuraa, että todennäköisyys sille, että osumia tulee vähintään yksi, on $1-e^{-\lambda}$. Tämä lauseke on teholaskujen kannalta hyvin tärkeä. (Vrt tappioprosenttilaskelmat sotilasaikakauslehdessä 1/51.) Laskelmien käytännöllinen suoritus ja tulosten vertailu yksinkertaistuvat ratkaisevasti puolilogaritmipaperia käyttämällä, koska työssä esiintyvät käyrät kuvautuvat sillä suorina.

e. Normaalijakautuma

Normaalijakautumaa voimme pitää vuorostaan Poissonin jakautuman sopivasti normeeraamalla saatuna rajamuotona, kun λ lähenee äärettömyyttä tai myös binomijakautuman rajamuotona, kun n lähenee äärettömyyttä p:n pysyessä vakiona. Emme esitä tässä mitään matemaattisia kaavoja, koska laskutoimituksia ei alkeellisin keinoin kuitenkaan voida suorittaa. Alempana normaalijakautuma tulee vielä monessa yhteydessä esiin. Normaalijakautuma on jatkuva x:n funktio, kun sensijaan edelliset ovat epäjatkuvia. Niissä voi muuttuja x saada vain täsmällisiä arvoja 0, 1, 2 . . ., mutta normaalijakautumassa ovat kaikki väliarvotkin mahdollisia. Keskiarvo, moodi ja keskeisarvo yhtyvät normaalijakautumassa.

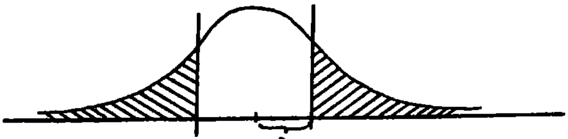
Kun erilaisia jakautumia kerrostuu päällekkäin, vaikuttaa yhdessä, on siitä nopeasti seurauksena normaalijakautuman muodostuminen, mikä sitten on hyvin yleinen jakautuma ja saanut siitä nimensäkin. Toisaalta voi tavallaan käydä päinvastoinkin. Esim yhdellä tykillä ammuttaessa voimme pitää pituus- ja sivupoikkeamien jakautumia normaaleina, mutta niiden yhdistelminä syntyvän sädepoikkeaman jakautuma ei ole "normaali" siinä mielessä, että yksiulotteista jakautumaa koskevat lait pätisivät sen suhteen, kuten usein otaksutaan.

Alla esitetään tavallisen 1-ulotteisen normaalijakautuman ominaisuuksista kaksiosainen taulukko

TAULUKKO 3

Normaalijakautuma taulukoituna

a) Normaalijakautuman prosenttipisteet



Odotusarvosta poikkeaman λ (standardipoikkeamina ja todennäköisinä poikkeamina) ulkopuolelle jäävä tapausten määrä = todennäköisyys sille, että poikkeama on suurempi kuin λ.

Ulkopuolelle jäävien %-luku	Poikkeama λ Stand.p.	Todennäk.p.
100	0,0000	0,0000
95	0,0627	0,0930
90	0,1257	0,1863
85	0,1891	0,2804
80	0,2533	0,3756
75	0,3186	0,4724
70	0,3853	0,5713
65	0,4538	0,6727
60	0,5244	0,7775
55	0,5978	0,8862
50	0,6745	1,0000
45	0,7554	1,1200
40	0,8416	1,2478
35	0,9346	1,3856
30	1,0364	1,5366

25	1,1503	1,7055
20	1,2816	1,9000
15	1,4395	2,1343
10	1,6449	2,4387
5	1,9600	2,9058
1	2,5758	3,8189
0,1	3,2905	4,8785
0,01	3,8906	5,7682

b) Vastaava taulukko kuin edellä mutta siten, että standardipoikkeamat ovat tasalukuja

Poikkeama λ		Ulkopuolelle jäävien
Stand.p.	Todennäk.p	%-luku
0,0	0,0000	100
0,2	0,2965	84,148
0,4	0,5930	68,916
0,6	0,8896	54,851
0,8	1,1861	42,371
1,0	1,4826	31,731
1,2	1,7791	23,014
1,4	2,0756	16,151
1,6	2,3722	10,960
1,8	2,6687	7,186
2,0	2,9652	4,550
2,2	3,2617	2,781
2,4	3,5582	1,640
2,6	3,8548	0,932
2,8	4,1513	0,511
3,0	4,4478	0,270
3,2	4,7443	0,137
3,4	5,0408	0,067
3,6	5,3374	0,032
3,8	5,6339	0,014
4,0	5,9304	0,006

(Arvot saatu Balltston taulukosta, todennäköisten poikkeamien arvot kuitenkin tekijän laskemat.)

f. Tasainen jakautuma

Tasainen jakautuma on jakautumista yksinkertaisin. Itse asiassa Poissonin jakautuma iskemätutkimuksissa edellyttää tasaista iskemäjakautumaa maastossa. Edellytyksenä ei ole, että jokaiseen yhtä suureen ruutuun maastossa tulisi juuri sama iskemätiheys kussakin ammunnassa, mutta todennäköisyys iskemälle tulee olla jokaisessa maaston kohdassa sama. Kun tutkitaan todennäköisyyksiä saada ruutuihin 0, 1, 2, ... iskemää, tullaan Poissonin jakautumaan (tai binomijakautumaan). Em tapauksessa maaston pinta-alaa on teoreettisesti pidettävä äärettömän suurena; käytännössä se on hyvinkin rajoitettu. Yleensä tasainen jakautuma alkaa (yksiulotteisena) jossakin pisteessä porrasmaisena ja loppuu täsmällisesti toisessa pisteessä, jolloin teoreettinen keskiarvo sattuu näiden pisteiden puoliväliin, samoin keskeisarvo. Moodi tietenkin puuttuu.

Esim. Mittaus on tehty 1 cm tarkkuudella, jolloin 0,5 cm alalle tasaisesti jakautuva keskimääräinen virhe on 0,25 cm. Keskivirhe on 0,29 cm, yleisesti a $\sqrt{3}/6$, kun a on tarkkuus (edellä 1 cm).

IV POIKKEAMAT JA VIRHEET

a. Yleistä

Kun esim ammumme yhdellä tykillä ampuma-arvoja muuttamatta, huomaamme, etteivät iskemät satu samaan paikkaan, vaan ammunnassa on todettavissa h a j o n t a a jonkin keskiarvon ympärillä. Poikkeamilla on kieltämättä syynsä, mutta näitä syitä on hyvin monta, emmekä pysty niitä hallitsemaan. Käytännöllisellä tarkkuudella kunkin laukauksen poikkeama (oikeasta) keskiarvosta on riippumaton muiden laukausten poikkeamista sikäli, ettemme voi esim edellisen poikkeaman perusteella ennustaa mitään seuraavan suhteen. Emme siis, ainakaan ilman erikoistoimenpiteitä, pysty korjaamaan ammuntaa millään tavoin pienentääksemme hajontaa. Olemme tekemisissä s a t u n n a i s m u u t t u j a n, stokastisen muuttujan kanssa. Tuo hajonnan satunnainen luonne ei kuitenkaan merkitse sitä, etteivät poikkeamat kokonaisuutena noudattaisi mitään lakia, päinvastoin on laskettavissa useitakin tunnuslukuja, joiden avulla hajonnan luonne tulee tarkemmin määritetyksi.

Kysymys satunnaisuudesta on lopultakin varsin suhteellinen käsite. Tiedämme ennakolta, että jokainen patteriston tykki ampuu systemaattisesti hieman eri tavalla, vaikka emme käytännössä kykene täysin hallitsemaan näitä eroja. Laskiessamme tilastollisessa mielessä näiden erojen vaikutusta esim. patteriston hajontaan (kyseessä siis suuri määrä patteristoja), voimme käsitellä niitä satunnaisluontoisina. Mutta tutkiessamme yhden tykin hajontaa, jää tällainen poikkeama systemaattisena ja hajontaan vaikuttamattomana pois. Tarkasti ottaen meidän on käsitettävä yhdelle ainoalle laukaukselleklin ominainen poikkeama, joka kaikissa muissa suhteissa on täysin satunnainen, juuri kyseisen laukauksen osalta systemaattiseksi. Edelleen tiettyä maalia ammuttaessa esiintyy vain sen suhteen systemaattisia poikkeamia, virheitä, jotka yleisesti, kaikkia maaleja ajatellen, ovat satunnaisia. On siis varottava jyrkästi pitämästä toisia poikkeamia tai virheitä aina satunnaisina, toisia systemaattisina, harkitsematta tarkemmin tehtävän asettelua.

Kysymys jakautumasta merkitsi samaa kuin satunnaissuuretta koskevien todennäköisyyksien laskeminen. Kun jakautuma on epäjatkuva, portaittain muuttuva, saavat todennäköisyydet pisteittäin nollasta eriäviä arvoja. Jatkuvassa jakautumassa on todennäköisyys jokaisessa pisteessä nolla, mutta ns kertymäfunktio kasvaa pisteestä toiseen siirryttäessä ja sen derivaatta, tiheysfunktio, antaa varsinaisesti kuvan todennäköisyyksistä tai esiintymien lukuisuudesta kunkin pisteen ympäristössä.

Hajontaa kuvaavina tunnuslukuina tulevat tavallisimmin kyseeseen

— keskimääräinen poikkeama eli poikkeamien (itseisarvojen) keskiarvo

— standardipoikkeama eli keskihajonta, myös neliökeskivirhe tai vain keskivirhe

— todennäköinen poikkeama tai virhe, myös neljännes- eli kvartiilipoikkeama.

Kaikki nämä poikkeamat voidaan laskea 1-, 2- tai 3-ulotteisia tapauksia varten. Tasossa ja avaruudessa voimme silloin puhua säde-poikkeamista ja niiden komponenteista eri ulottuvuuksissa. Keskiarvon tai muidenkin keskilukujen poikkeamia nimitämme tarkoituksenmukaisuussyistä tavallisimmin virheiksi.

b. Poikkeamat ja niiden suhdeluvut

Tehtävä on joskus sen laatuinen, että voimme ennakolta, ilman havaintoja ja kokeita, laskea em poikkeamien numeroarvot. Toisinaan tunnemme tarkasteltavan suureen tarkan arvon, mutta emme ole selvillä sen hajonnasta, mikä näin ollen on mitattava ja laskettava. Ehkä tavallisimmin olemme tilanteessa, jolloin emme tunne suureen oikeata arvoa emmekä hajonnan ominaisuuksia. Laskelmat muodostuvat kussakin tapauksessa erilaisiksi.

Esimerkkinä tapauksesta, jolloin ennakolta tunnemme laskuissa tarpeelliset arvot, on osumatodennäköisyyksien laskeminen Poissonin jakautuman mukaan suurehkon aluemaalin sisällä oleviin korsuihin ym rakenteisiin, joiden haavoittuva pinta-ala on tunnettu. Laukaus (iskemä-) luku n, aluemaalin suuruus A ja rakenteiden haavoittuva pinta-ala a määrittävät keskimääräisen iskemäluvun $\lambda = \frac{na}{A}$ maalia kohti, jolloin jakautuma (0, 1, 2 ... osumaa) on suoraan laskettavissa, samoin standardipoikkeama $= \sqrt{\lambda}$.

Kun tutkimme patterien tai patteristojen yleistä osumatarkkuutta pistemaaliin tietyllä ampumaetäisyydellä, on meillä jokaisessa ammunnassa maalin paikka ennakolta tiedossa, mutta poikkeaman suhteen on tyydyttävä likimääräiseen määrittelyyn, sillä oikeata iskemäkeskipistettä emme saa aivan tarkasti selville. Mittaamalla laukausten poikkeamat maalista ja ottamalla keskiarvo joka ammunnassa, on osumatarkkuuden tunnusluvut likimäärin laskettavissa. Ollaksemme tarkkoja on iskemäkeskipisteen määrittämisvirhe otettava tällöin vähennyksenä huomioon siten, että maalipoikkeamien varianssista (standardipoikkeaman neliö) vähennetään iskemäkeskipisteen (määritys-) virheen varianssi. Vastaava tilanne voi muodostua myös silloin, kun poikkeamien mittaluvut jäävät epätarkoiksi puutteellisten mittausmenetelmien vuoksi.

Jos edelleen tutkimme ammuntatekijöitä siinä mielessä, että haluamme laskea hajonta-arvot (oikean) iskemäkeskipisteen suhteen, joudumme myös pulman eteen siinä, ettemme saa tarkoin selville iskemäkeskipisteen sijaintia. Laukausten hajonta on laskettava iskemäkuvion painopisteestä, mikä kuitenkaan ei tarkoin yhdy oikeaan iskemäkeskipisteeseen. Tämä seikka on otettava ja voidaan ottaa tarkkuuteen pyrittäessä laskuissa huomioon.

Käytämme seuraavia merkintöjä

k = keskimääräinen poikkeama

s = standardipoikkeama (etenkin koko perusjoukon, siis oikeasta ja tarkasta poikkeamasta on vakiintunut merkintä σ, sigma)

t = todennäköinen poikkeama (nimitys siviilikäytössä vanhentunut, mutta kannattanee säilyttää sotilasterminä)

Tasossa liitämme alaviitan r, joten saamme

k_r = keskimääräinen sädepoikkeama

s_r = standardisädepoikkeama

t_r = todennäköinen sädepoikkeama

Avaruudessa käytämme alaviittaa R, jolloin vastaavat poikkeamat ovat k_R, s_R ja t_R.

Keskimääräinen poikkeama lasketaan keskiarvosta, ellei oikea suureen arvo ole tunnettu. (Jos se erikoistapauksessa lasketaan keskeisarvosta, saadaan sille p i e n i n m a h d o l l i n e n a r v o.) Kun poikkeamaa yleisesti merkitsemme p_i, on keskimääräinen poikkeama tapausluvulla n

$$k = \frac{|p_1| + \ldots + |p_n|}{n}$$

Jos standardipoikkeama voidaan laskea suureen oikeasta keskiarvosta (odotusarvosta) tarkasti mitattuja poikkeamia käyttäen, se on

$$s = \sqrt{\frac{p_1^2 + \ldots + p_n^2}{n}} \quad (=\sigma)$$

tai suureen havaintoarvoista $x_1, x_2, \ldots, x_n$ suoraan laskettuna, kun oikea arvo (odotusarvo) on x_o

$$s = \sqrt{\frac{x_1^2 + \ldots + x_n^2}{n} - x_o^2}$$

Kun n on suuri, voidaan menetellä näin, vaikka (virheettömät) mittaukset onkin tehty havainnoista saadusta keskiarvosta. Sama koskee käytännössä tapauksia, jolloin poikkeamien mittaaminen tai muu määrittäminen on suoritettu epätarkasti, eikä senvuoksi pyritäkään suureen lopputarkkuuteen.

Todennäköinen poikkeama t on suuruudeltaan sellainen, että sitä pienempiä ja suurempia poikkeamia esiintyy yhtä paljon. Usein se lasketaan johtamalla muista poikkeamista (muuntamiskertoimilla).

Jos mittausten lähtökohtana on havaittu (siis likimääräinen) keskiarvo, on em tavalla saadut poikkeamat periaatteessa (normaalijakautumassa) kerrottava juurilausekkeella $\sqrt{\frac{n}{n-1}}$, joten esim standardipoikkeaman tarkempi arvo silloin on

$$s = \sqrt{\frac{p_1^2 + \ldots + p_n^2}{n-1}}$$

tai havaintoarvoista suoraan laskettuna, kun havaittu keskiarvo on $\bar{x}$

$$s = \sqrt{\frac{x_1^2 + \ldots + x_n^2}{n-1} - \frac{n}{n-1} \cdot \bar{x}^2}$$

[Näin lasketun standardipoikkeaman jakautuma on käytännöllisellä tarkkuudella normaali, kun n (tarkemmin: vapausasteiden luku) nousee yli 30.]

Tarkempi laskutapa antaa suuremman lopputuloksen. Mittausten epätarkkuus vaikuttaa samalla tavalla, joten usein on siis aiheetonta käyttää tarkempaa kaavaa.

Oikeasta arvosta mitattujen poikkeamien avulla laskettu standardipoikkeama (juuren alla n) ja keskiarvosta tarkalla tavalla laskettu poikkeama (juuren alla n—1) ovat samanarvoisia, ellei n ole aivan pieni luku.

Keskiarvon standardipoikkeama eli keskivirhe on

$$s_0 = \frac{s}{\sqrt{n}} = \frac{\sqrt{p_1^2 + \ldots + p_n^2}}{n}$$

tai hieman tarkemmin (kerrointa $\sqrt{\frac{n}{n-1}}$ käyttäen) koska se lasketaan tietenkin havaintoihin perustuen.

$$s_0 = \sqrt{\frac{p_1^2 + \ldots + p_n^2}{n\,(n-1)}}$$

(Erikoisuutena, josta joskus on hyötyä mainittakoon, että tiheysfunktion kuvaajassa on käännepiste yhden standardipoikkeaman päässä keskustasta.)

Suhdeluvut keskeisarvon keskivirheen laskemista varten on esitetty edellä taulukossa 1 ja hajonta-alueen puolittamisella saadun hajontakeskuksen keskivirheen suhdeluvut taulukossa 2.

Siinä yksinkertaisessa tapauksessa, että poikkeamat k ovat eri akselien suunnissa yhtä suuret ($k_x = k_y = k_z$) ja että ne ovat vain radiusvektorin funktioita (normaalijakautuma ei välttämätön), saadaan seuraavat kiinteät suhdeluvut

$$k = \frac{2}{\pi}\,k_r = 0{,}6366\,k_r = 0{,}5\,k_R$$

$$k_r = \frac{\pi}{2}\,k = 1{,}5708\,k = \frac{\pi}{4}\,k_R = 0{,}7854\,k_R$$

$$k_R = 2\,k = \frac{4}{\pi}k_x = 1{,}2732\,k_x$$

Esim 1. Heittimellä ammuttaessa havaitaan, että pituus- ja sivupoikkeamat ovat suunnilleen yhtä suuret, kumpikin keskimäärin = a. Sädevirhe k_r on silloin keskimäärin = 1,57 a.

Esim 2. Keskituuli vaihtelee puuskaisuuden takia siten, että puuskakomponentit kaikkien kolmen akselin suunnissa ovat suunnilleen 1 m/sek. Horisontaalinen sädepoikkeama (k_r) keskituulesta on silloin 1.57 m/sek ja kokonaispoikkeama (k_R) 2 m/sek, jolloin siis vertikaalinenkin heilahtelu on otettu mukaan.

Varsin tavallinen on tapaus, jolloin akselien suunnissa vallitsee normaalijakautuma, mutta keskimääräiset poikkeamat k_x, k_y ja k_z ovat eri suuret. Ellei suurimman ja pienimmän keskimääräisen poikkeaman suhde ole kovin suuri (korkeintaan 2), voimme tyytyä seuraaviin likimääräiskaavoihin

$$k_r = 1.57 \sqrt{k_x \cdot k_y}$$

$$k_B = 2 \sqrt[3]{k_x \cdot k_y \cdot k_z}$$

Standardipoikkeamien keskinäiset suhteet ovat hyvin yksinkertaiset. Jakaantumislaista riippumatta saadaan, jos eri akselien suuntaisia standardipoikkeamia merkitään s_x, s_y ja s_z:

$$s_r = \sqrt{s_x^2 + s_y^2}$$

$$s_B = \sqrt{s_x^2 + s_y^2 + s_z^2}$$

Jos akselien suuntaiset poikkeamat ovat yhtäsuuret $= s$, supistuvat kaavat seuraaviksi:

$$s_r = \sqrt{2}\, s = 1.4142\ s$$

$$s_B = \sqrt{3}\, s = 1.7321\ s$$

Todennäköisten poikkeamien keskinäiset suhteet riippuvat jakautumalaista. Jos jakautuma on normaali ja sama kaikkien akselien suunnissa ($= t$), on 50 % kaikista tapauksista sisältävän todennäköisyysympyrän säde

$$t_r = 1.7456\ t$$

Jos taas todennäköiset poikkeamat akselien suunnissa ovat erilaiset, t_x, t_y ja t_z, ovat vastaavan todennäköisyysellipsin puoliakselit

$$a = 1.7456\ t_x$$
$$b = 1.7456\ t_y$$

ja ellipsoidissa vastaavasti lisäksi

$$c = 1.7456\ t_z$$

Mikäli korvaamme ellipsin ympyrällä, saamme likimäärin

$$t_r = 1.75 \sqrt{t_x \cdot t_y}$$

50 %:n todennäköisyyspallon säde saman normaalijakautuman vallitessa akselien suunnissa (todennäk. poikkeama = t) on

$$t_R = 2.2805\ t$$

Jos jakautumat akselien suunnissa ovat erilaiset, voidaan kyseinen todennäköisyysellipsoidi korvata pallolla vastaavasti kuin tasossa ellipsi ($t_R = 2.28 \sqrt[3]{t_x \cdot t_y \cdot t_z}$).

Edellä on verrattu toisiinsa saman lajisia poikkeamia. On vielä tarpeen rinnastaa keskenään eri laatuiset poikkeamat. Jos jakautuma on normaali ja sama akselien suunnissa, saadaan

$$k = \sqrt{\frac{2}{\pi}}\ s = 0{,}7979\ s$$

$$k_r = \sqrt{\frac{\pi}{2}}\ s = 1.2533\ s$$

$$k_R = 2\sqrt{\frac{2}{\pi}}\ s = 1.5958\ s$$

Voimme käyttää myös seuraavia likiarvoja

$$k_r = 1.25 \sqrt{s_x \cdot s_y}$$

$$k_R = 1.6 \sqrt[3]{s_x \cdot s_y \cdot s_z}$$

jolloin jälleen suurin akselin suuntainen keskipoikkeama ei saisi pienempään verrattuna olla suurempi kuin kaksinkertainen.

Vastaavasti saamme, kun jakautuma on akselien suunnissa sama ja normaali

$$t = 0.6745\ s$$

$$t_r = 1.1774\ s$$

$$t_R = 1.5382\ s$$

tai poikkeamien akselien suunnissa ollessa erilaiset

$$t_r = 1.18 \sqrt{s_x \cdot s_y}$$

$$t_R = 1.54 \sqrt[3]{s_x \cdot s_y \cdot s_z}$$

Edelleen saamme

$$k = 1.1830 \; t$$

$$k_r = 1.8581 \; t = 1.0645 \; t_r$$

$$k_R = 2.3659 \; t = 1.0374 \; t_R$$

kun akselin suuntaisilla jakautumilla ei ole eroa, tai likiarvoisesti, erojen ollessa vähäisiä

$$k_r = 1.86 \sqrt{t_x \cdot t_y}$$

$$k_R = 2.37 \sqrt[3]{t_x \cdot t_y \cdot t_z}$$

Esim. Kiväärillä 100 m matkalla ammuttaessa on saatu todennäköisiksi poikkeamiksi pysty- ja vaakasuunnissa 3 cm. Vastaava keskimääräinen poikkeama on silloin $k = 1{,}18 \cdot 3 = 3{,}5$ cm, todennäköinen sädepoikkeama $t_r = 1{,}75 \cdot 3 = 5{,}25$ cm, keskimääräinen sädepoikkeama $k_r = 1{,}86 \cdot 3 = 5{,}6$ cm. Standardipoikkeamat ovat $s = 3/0.6745 = 4{,}4$ cm, $s_r = \sqrt{2}\, s = 6{,}25$ cm.

Esitämme vielä seuraavan yhteenvetotaulukon:

TAULUKKO 4

Keskimääräisten, todennäköisten ja standardipoikkeamien keskinäiset muuntamiskertoimet normaalijakautumassa ja akselien suuntaisten komponenttien ollessa yhtä suuret

$$
\begin{array}{llll}
k = & & = 0.6366 \; k_r & = 0.5000 \; k_R \\
& = 1.1830 \; t & = 0.6777 \; t_r & = 0.5187 \; t_R \\
& = 0.7979 \; s & = 0.5642 \; s_r & = 0.4607 \; s_R \\
k_r = & 1.5708 \; k & & = 0.7854 \; k_R \\
& = 1.8581 \; t & = 1.0645 \; t_r & = 0.8148 \; t_R \\
& = 1.2533 \; s & = 0.8862 \; s_r & = 0.7236 \; s_R
\end{array}
$$

k_R	$= 2.0000\ k$	$= 1.2732\ k_r$	
	$= 2.3659\ t$	$= 1.3553\ t_r$	$= 1.0374\ t_R$
	$= 1.5958\ s$	$= 1.1284\ s_r$	$= 0.9213\ s_R$
t	$= 0.8453\ k$	$= 0.5382\ k_r$	$= 0.4226\ k_R$
		$= 0.5729\ t_r$	$= 0.4385\ t_R$
	$= 0.6745\ s$	$= 0.4769\ s_r$	$= 0.3894\ s_R$
t_r	$= 1.4756\ k$	$= 0.9394\ k_r$	$= 0.7378\ k_R$
	$= 1.7456\ t$		$= 0.7654\ t_R$
	$= 1.1774\ s$	$= 0.8326\ s_r$	$= 0.6798\ s_R$
t_R	$= 1.9278\ k$	$= 1.2273\ k_r$	$= 0.9639\ k_R$
	$= 2.2805\ t$	$= 1.3064\ t_r$	
	$= 1.5382\ s$	$= 1.0877\ s_r$	$= 0.8881\ s_R$
s	$= 1.2533\ k$	$= 0.7979\ k_r$	$= 0.6267\ k_R$
	$= 1.4826\ t$	$= 0.8493\ t_r$	$= 0.6501\ t_R$
		$= 0.7071\ s_r$	$= 0.5774\ s_R$
s_r	$= 1.7724\ k$	$= 1.1284\ k_r$	$= 0.8863\ k_R$
	$= 2.0967\ t$	$= 1.2011\ t_r$	$= 0.9194\ t_R$
	$= 1.4142\ s$		$= 0.8165\ s_R$
s_R	$= 2.1708\ k$	$= 1.3820\ k_r$	$= 1.0855\ k_R$
	$= 2.5679\ t$	$= 1.4700\ t_r$	$= 1.1260\ t_R$
	$= 1.7321\ s$	$= 1.2247\ s_r$	

(Taulukko tekijän laatima)

c. Hajonta-alueen pituus

Hajonta-alueen pituus (esim ammunnassa pisimmän ja lyhimmän iskemän väli) antaa nopeasti likimääräisen käsityksen hajonnan suuruudesta. Alla esitetään normaalijakautumaa koskeva taulukko, jossa hajonta-alueen keskimääräinen pituus ilmaistaan standardipoikkeamina ja todennäköisinä poikkeamina tapausluvusta n (esim iskemäluku) riippuvana. Hajonta-alueen pituuksien jakautuma on vino: moodi on keskimääräistä arvoa pienempi. Koska arvot ovat horjuvia, esitetään tapauslukuun 20 asti myös ne pituuksien arvot standardipoikkeamina, joita satunnaisista syistä ei 99 %:n varmuudella ylitetä tai myöskään aliteta. Suuremmille tapausluvuille kuin 20 ei viimeksi mainittuja arvoja yleensä esitetä, koska se saattaisi olla harhaan johtavaa. Pienetkin häiriöt normaalijakautumasta (mitä ammunnoistakin puheen ollen sattuu) pilaavat tällöin heti tarkkaa normaalijakautumaa koskevat tulokset, koska ne riippuvat kokonaan kaukana hajonnan keskuksesta olevista reuna-arvoista.

Taulukko 5

Hajonta-alueen keskimääräinen pituus

Tapaus-luku n	Keskim. pituus		99 %:n varmuudella	
	stand.p.	tod.näk.p.	vähintään stand.p.	enintään stand.p.
2	1.128	1.673	0.02	3.64
3	1.693	2.509	0.19	4.12
4	2.059	3.052	0.43	4.40
5	2.326	3.448	0.66	4.60
6	2.534	3.758	0.87	4.76
7	2.704	4.009	1.05	4.88
8	2.847	4.221	1.20	4.99
9	2.970	4.403	1.34	5.08
10	3.078	4.563	1.47	5.16
11	3.173	4.704	1.58	5.23
12	3.258	4.831	1.68	5.29
13	3.336	4.946	1.77	5.35
14	3.407	5.051	1.86	5.40
15	3.472	5.147	1.93	5.45
16	3.532	5.237	2.01	5.49
17	3.588	5.319	2.07	5.54
18	3.640	5.397	2.14	5.57
19	3.689	5.469	2.20	5.61
20	3.735	5.537	2.25	5.65
21	3.778	5.602		
22	3.819	5.663		
23	3.858	5.720		
24	3.895	5.775		
25	3.931	5.828		
30	4.086	6.057		
35	4.213	6.247		
40	4.322	6.407		
45	4.415	6.546		
50	4.498	6.669		
55	4.572	6.778		
60	4.639	6.877		
65	4.699	6.967		
70	4.755	7.049		
75	4.806	7.125		
80	4.854	7.196		
85	4.898	7.262		
90	4.939	7.323		
95	4.978	7.381		
100	5.015	7.436		
150	5.298	7.856		
200	5.492	8.143		
300	5.756	8.533		
400	5.936	8.801		
500	6.073	9.004		
1000	6.483	9.612		

(Keskim. pituudet stand.poikkeamina LHC Tippett, Biometrika 17, 1925 ss 386—387, tekijä laskenut tod.näk.poikkeamat. Viimeiset sarakkeet A Hald: Statistical Tables and Formulas, 1952 ss 60—61.)

Esim 1. Tykillä ammutaan 10 laukausta ennakolta tunnetun todennäköisen poikkeaman ollessa 20 m. Jos koe suoritetaan useampaan kertaan ja mitataan hajonta-alueen pituus joka kerta erikseen, on odotettavissa oleva keskiarvo 4.56 × 20 m = 91 m ja vain yhdessä tapauksessa sadasta jää joku yksittäisarvo pienemmäksi kuin 1.48 × 1.47 × 20 m = 43.5 m ja yhtä harvoin sattuu suurempia arvoja kuin 1.48 × 5.16 × 20 m = 153 m (kerroin 1.48 otettu taulukosta 4).

Esim 2. Tiede ja Ase N:o 13:ssa vuodelta 1955 sivulla 252 on tekijä keinotekoisiin koesarjoihin nojautuen esittänyt, että tykin hajonta-alueen käytännöllinen minimipituus on 25 % ja maksimipituus 225 % keskimääräisestä arvosta, patterin vastaavasti 60 % ja 170 % sekä patteriston 70 % ja 150 %. Kysymyksessä on iskun ampuminen, jolloin vastaavat laukausluvut ovat 6, 24 ja 72. Edelleen on todettu, että paitsi tykin myös patterin hajonnan jakautumaa voidaan pitää kutakuinkin normaalina, patteriston hajonnan sensijaan jossakin määrin poiketessa normaalista. Kontrolloimme näillä edellytyksillä tuloksen ensin tykin osalta. Taulukossa 5 vastaa tapauslukua 6 keskimääräinen arvo 2.534, josta 25 % tekee 0.633. Tämä jää alle taulukon "vähintään"-sarakkeessa ilmoitetun luvun 0.87, joten sen todennäköisyysluku on alle 1 prosentin. Tarkempi graafinen interpolointi alkuperäisestä taulukosta antaa tuloksen 0,2 %. Siis esitetty käytännöllinen minimi alitetaan 2 tapauksessa tuhannesta. 225 % keskiarvosta taas tekee 5.702 ja "enintään"-sarakkeesta löydämme luvun 4.76. Tarkempi vertailu jälleen osoittaa, että esitetty maksimi (mikä on ammunnasta puheenollen vaarallisempi) ylitetään harvemmin kuin yhdessä tapauksessa tuhannesta. Patterin osalta on alkuperäisessä taulukossakin turvauduttava interpoloinnin ohella myös ekstrapolointiin. Minimiä ei aliteta useammin kuin 5 tapauksessa tuhannesta ja maksimi ylitetään harvemmin kuin yhdessä tapauksessa tuhannesta. Patteriston osalta ei tarkistusta voida suorittaa.

KÄYTETTYJÄ LÄHTEITÄ:

Tekn korkeakoulun luentosarjoja (P J Myrberg, O Lokki)

Biometrika, eri vuosikertoja

G Elfving: Todennäköisyyslaskenta. 1956.

Fisher and Yates: Statistical Tables. 1957.

A Hald: Statistical Tables and Formulas. 1952.

Alameri—Pöyhönen: Johdatusta tilastolliseen tutkimukseen. 1954.

The Canadian Surveyor, 1961 numero 1.

J M Juran: Quality—Control Handbook. 1951.

Berry—Bollay—Beers: Handbook of Meteorology. 1945.

Operations Research (ORSA) 1/1961.

Pohjoismainen tilastosanasto. 1954.