Aki Taanila

TODENNÄKÖISYYSLASKENTAA JA -JAKAUMIA

17.4.2018

SISÄLLYS

JOHDANTO

Satunnaisilmiöt ovat ilmiöitä, joiden lopputulokseen sattuma vaikuttaa. Tällaisten ilmiöiden lopputulosta ei voida varmuudella ennustaa. Sen sijaan monien satunnaisilmiöiden kohdalla voidaan selvittää todennäköisyydet eri lopputuloksille. Joillekin satunnaisilmiöille lopputulosten todennäköisyydet ovat tarkkaan laskettavissa (esim. ruletti ja monet muut uhkapelit). Joillekin satunnaisilmiöille taas voidaan arvioida eri lopputulosten todennäköisyydet historiatiedoista (tilastollinen todennäköisyys). Monien satunnaisilmiöiden voidaan olettaa noudattavan likimain jotain mallia (todennäköisyysjakaumaa).

Tässä monisteessa käsitellään todennäköisyyksiin liittyviä laskusääntöjä ja satunnaisilmiöiden mallintamista todennäköisyysjakaumien avulla. Mukana ovat tarvittavat Excel-ohjeet.

Viimeisin versio tästä monisteesta ja siihen liittyvästä materiaalista löytyy osoitteesta

https://tilastoapu.wordpress.com/mallinna/

Myös monisteessa viitatut Excel-esimerkit löytyvät yllä mainitusta osoitteesta.

Huomautus Excel 2010 -käyttäjille: Joidenkin funktioiden nimet on uusittu. Vanhat nimet kuitenkin toimivat edelleen ja niitä pitää käyttää, jos tarvitset yhteensopivuutta aikaisempiin versioihin.

BINOMDIST (BINOMIJAKAUMA) uusi nimi **BINOM.DIST** (BINOMI.JAKAUMA).

POISSON (POISSON) uusi nimi **POISSON.DIST** (POISSON.JAKAUMA).

NORMDIST (NORM.JAKAUMA) uusi nimi **NORM.DIST** (NORMAALI.JAKAUMA).

NORMINV (NORM.JAKUAMA.KÄÄNT) uusi nimi **NORM.INV** (NORMAALI.JAKAUMA.KÄÄNT).

EXPONDIST (EKSPONENTIAALIJAKAUMA) uusi nimi **EXPON.DIST** (EKSPONENTIAALI.JAKAUMA).

Tutustu myös muihin oppimateriaaleihini http://myy.haaga-helia.fi/~taaak/

KOMBINAATIO-OPPIA

Kombinaatio-opissa lasketaan tulosvaihtoehtojen lukumääriä, esimerkiksi

- kuinka monella tavalla lottoruudukko voidaan täyttää
- kuinka monta kolme kirjainta ja kolme numeroa sisältävää rekisterikilpeä voidaan muodostaa
- kuinka monella tavalla 9 henkilöä voidaan jakaa 3 hengen tiimeihin
- kuinka moneen eri järjestykseen korttipakan kortit voidaan sekoittaa.

Kombinaatio-opin tehtävät voidaan ratkaista käyttämällä neljää laskusääntöä: tuloperiaate, järjestysten lukumäärä, variaatio ja kombinaatio.

Tuloperiaate

Jos suoritetaan k peräkkäistä toisistaan riippumatonta valintaa, joista ensimmäinen voidaan tehdä n_1:lla tavalla, toinen n_2:lla tavalla jne., niin erilaisia valintojen yhdistelmiä on $\mathbf{n_1 \cdot n_2 \cdot \ldots \cdot n_k}$ kappaletta.

Esimerkki. Kuinka monta erilaista veikkausriviä on olemassa? Oletetaan, että veikattavana on 13 kohdetta, joista kussakin on mahdollista valita 1, x tai 2. Kyseessä on siis 13 peräkkäistä valintaa, joista kussakin on 3 vaihtoehtoa. Erilaisia valintojen yhdistelmiä on tuloperiaatteen mukaisesti 3^{13} = 1 594 323 kappaletta.

Esimerkki. Pizzabaari tarjoilee jauheliha, tonnikala, kinkku ja kasvispizzoja. Mausteeksi voi valita pippuria, sipulia, valkosipulia, tomaattisosetta ja oliiveja. Kuinka monta erilaista pizzaa on valittavana?

Esimerkki. Oletetaan, että pizzan tekee erilaiseksi erilainen täyte, mutta myös erilaiset mausteet. Voidaan ajatella, että kyseessä on peräkkäisiä valintoja (valitaan täyte neljän vaihtoehdon joukosta, valitaan otetaanko pippuria vai ei, valitaan otetaanko sipulia vai ei jne.). Tuloperiaatteen mukaan valintojen yhdistelmiä on $4 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 128$ kappaletta.

Järjestysten lukumäärä

Jos joukossa on n-alkioita, niin alkiot voidaan järjestää **n!** (lue: n kertoma) erilaiseen järjestykseen.

Tässä on kyse tuloperiaatteen soveltamisesta. Järjestäminen voidaan nimittäin ajatella peräkkäisiksi valinnoiksi. Järjestetyn joukon ensimmäinen alkio voidaan valita n tavalla, toinen alkio n-1 tavalla, kolmas alkio n-2 tavalla jne. Tuloperiaatteen mukaan erilaisia vaihtoehtoja on $n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 1 = n!$ kappaletta.

Esimerkki. Kirjaimet A, B ja C voidaan järjestää $3! = 3 \cdot 2 \cdot 1 = 6$ tavalla. Tämän voi helposti todentaa kokeilemalla (ABC, ACB, BAC, BCA, CAB, CBA).

Esimerkki. 8 ravihevosta voi saapua maaliin 8! = 40 320 eri järjestyksessä.

Esimerkki. 10 alokasta voi asettua riviin 10! = 3 628 800 eri järjestyksessä.

Esimerkki. Korttipakan kortit (52 kpl) voidaan sekoittaa $52! \approx 8{,}066 \cdot 10^{67}$ eri järjestykseen.

Vertailun vuoksi: Jos maailmankaikkeuden iäksi oletetaan 15 miljardia vuotta, niin sekunteja maailmankaikkeuden alusta on kulunut noin $4{,}7 \cdot 10^{17}$ sekuntia. Jos maailmankaikkeuden alkuhetkestä alkaen olisi sekoitettu korttipakka joka sekunti uuteen järjestykseen, niin toistaiseksi ei vielä olisi ehditty käydä kaikkia mahdollisia järjestyksiä läpi.

Excelissä n! lasketaan funktiolla **=FACT(n)** (KERTOMA).

Variaatio

Jos n alkiota sisältävästä joukosta poimitaan k alkiota siten, että poimittujen alkioiden järjestys huomioidaan, niin erilaisia mahdollisuuksia on $\frac{n!}{(n-k)!}$ kappaletta.

Olennaista on, että sama k alkiota sisältävä joukko lasketaan mukaan useampaan kertaan, koska erilaiset järjestyksetkin huomioidaan.

Variaatioissakin on kyse tuloperiaatteen soveltamisesta. Järjestetyn joukon poimiminen voidaan ajatella peräkkäisinä valintoina. Ensimmäinen alkio voidaan poimia n tavalla, toinen n-1 tavalla jne. Viimeinen poimittava alkio voidaan poimia n-k+1 tavalla. Laskemalla tuloperiaatteen mukaisesti $n \cdot (n-1) \cdot \ldots \cdot (n-k+1)$ päädytään samaan tulokseen kuin variaatio-kaavaa käyttäen.

Esimerkki. Kirjainten A, B ja C joukosta voidaan valita kahden kirjaimen järjestettyjä joukkoja

$$\frac{3!}{(3-2)!} = 3 \cdot 2 = 6 \text{ kappaletta}$$

Tämä voidaan todentaa kokeilemalla (AB, AC, BA, BC, CA, CB).

Esimerkki. 8 ravihevosta voi jakaa 3 ensimmäistä sijaa

$$\frac{8!}{(8-3)!} = 336 \text{ tavalla}$$

Excelissä voit laskea variaatioiden määrän funktiolla **=PERMUT(n;k)** (PERMUTAATIO).

Kombinaatio

Jos n alkiota sisältävästä joukosta poimitaan k alkiota siten, että poimittujen alkioiden järjestystä ei huomioida, niin erilaisia mahdollisuuksia on $\frac{n!}{k!(n-k)!}$ kappaletta.

Toisin kuin variaatioissa, kombinaatioissa kukin k alkiota sisältävä joukko lasketaan mukaan vain kertaalleen, koska erilaisia järjestyksiä ei huomioida.

Kombinaatioiden laskukaava muistuttaa variaatioiden laskukaavaa. Ainoana erona on jakajana oleva k!, joka poistaa erilaiset järjestykset.

Esimerkki. Lotossa 39 numeron joukosta valitaan 7 numeroa. Lottoruudukkoa täytettäessä ei ole väliä, missä järjestyksessä 7 numeroa valitaan. Olennaista on vain se, mitkä 7 numeroa valitaan. Kombinaatioiden laskukaavalla saadaan mahdollisten lottoruudukoiden lukumääräksi

$$\frac{39!}{7!(39-7)!} = 15380937$$

Kuinka monta erilaista viiden kortin kättä (järjestystä ei huomioida) voidaan 52 kortin pakasta nostaa?

$$\frac{52!}{5!(52-5)!} = 2598960$$

Kuinka moni kaikista mahdollisista viiden kortin käsistä ei sisällä yhtään ässää? Valitaan 5 korttia 48 kortin joukosta, koska ässiä ei saa olla mukana:

$$\frac{48!}{5!(48-5)!} = 1712304$$

Kuinka moni kaikista mahdollisista viiden kortin käsistä sisältää tasan yhden ässän? Valitaan 4 korttia 48 kortin joukosta ja tämän jälkeen valitaan yksi ässä neljän mahdollisen ässän joukosta. Tuloperiaatetta käyttäen saadaan:

$$\frac{48!}{4!(48-4!)} \cdot 4 = 778320$$

Kuinka moni kaikista mahdollisista viiden kortin käsistä sisältää tasan kaksi ässää? Valitaan 3 korttia 48 kortin joukosta ja tämän jälkeen kaksi ässää neljän mahdollisen ässän joukosta. Tuloperiaatetta käyttäen saadaan:

$$\frac{48!}{3!(48-3)!} \cdot \frac{4!}{2!(4-2)!} = 103776$$

Kuinka moni kaikista mahdollisista viiden kortin käsistä sisältää tasan kolme ässää? Valitaan 2 korttia 48 kortin joukosta ja tämän jälkeen kolme ässää neljän mahdollisen ässän joukosta. Tuloperiaatetta käyttäen saadaan:

$$\frac{48!}{2!(48-2)!} \cdot \frac{4!}{3!(4-3)!} = 4512$$

Kuinka moni kaikista mahdollisista viiden kortin käsistä sisältää neljä ässää? 4 ässää ei voi tulla kuin yhdellä tavalla ja niiden kaverina voi olla mikä tahansa 48 kortista, joten mahdollisuuksia on 48 kappaletta.

Edellä tarkasteltujen tapausten pitäisi kattaa kaikki mahdolliset viiden kortin kädet. Tarkistetaan lopuksi, että näin on:

1 712 304+778 320+103 776+4 512+48=2 598 960

Tulos on sama kuin aiemmin laskettu viiden kortin käsien lukumäärä.

Excelissä voit laskea kombinaatioiden määrän funktiolla **=COMBIN(n;k)** (KOMBINAATIO).

TODENNÄKÖISYYDEN LASKUSÄÄNTÖJÄ

Jatkossa tarvitaan seuraavia käsitteitä ja määritelmiä:

- Alkeistapahtuma = yksittäinen tulosmahdollisuus.
- Tapahtuma = joukko alkeistapahtumia (esimerkiksi parillinen silmäluku nopanheitossa).
- Tapahtumia merkitään isoilla aakkosilla A, B, jne.
- Tapahtuman todennäköisyys on tapahtumalle suotuisien alkeistapahtumien määrä jaettuna kaikkien alkeistapahtumien määrällä (esimerkiksi parillisen silmäluvun todennäköisyys nopanheitossa 3/6=0,5).
- Tapahtuman A todennäköisyyttä merkitään P(A).
- Erillisyys: Tapahtumat A ja B ovat erillisiä, jos ne eivät voi sattua yhtä aikaa. Esimerkiksi tapahtumat ’nostettu kortti on pata’ ja ’nostettu kortti on hertta’ ovat erillisiä. Sen sijaan tapahtumat ’nostettu kortti on kuvakortti’ ja ’nostettu kortti on hertta’ eivät ole erillisiä, koska nostettu kortti voi olla yhtä aikaa kuvakortti ja hertta.
- Riippumattomuus: Tapahtumat A ja B ovat toisistaan riippumattomia, jos toisen tapahtuminen ei vaikuta toisen todennäköisyyteen. Esimerkiksi tapahtumat 'ensimmäinen nostettu kortti on ässä' ja 'toinen nostettu kortti on ässä' eivät ole riippumattomia. Toisen kortin kohdalla ässän todennäköisyys riippuu siitä, mikä ensimmäinen kortti oli.

Yhteenlaskusääntö

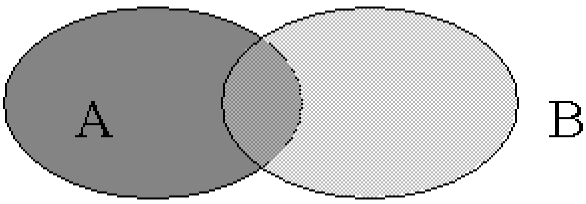
Todennäköisyys, että sattuu tapahtuma A tai tapahtuma B (jompikumpi tai molemmat):

P(A tai B) = P(A) + P(B) – P(A ja B)

Jos tapahtumat A ja B ovat erilliset, niin

P(A tai B) = P(A) + P(B)

Yhteenlaskusääntöä voidaan havainnollistaa kuvaamalla tapahtumia A ja B joukkoina ja samaistamalla joukkojen pinta-alat todennäköisyyksien kanssa.



Jos molemmat tapahtumista A ja B voivat sattua yhtä aikaa, niin joukoilla on yhteistä aluetta. Laskettaessa A:n ja B:n todennäköisyydet yhteen, tulee yhteinen alue lasketuksi kahteen kertaan. Niinpä yhteinen alue täytyy vähentää todennäköisyyksien summasta. Tapahtumat voivat olla myös erillisiä, jolloin yhteistä aluetta ei ole.

Tuotteen laatu vaihtelee pitkäaikaisten tilastojen mukaan seuraavasti:

- materiaalivika A todennäköisyydellä 0,1
- käsittelyvika B todennäköisyydellä 0,2
- molemmat viat todennäköisyydellä 0,05 (yhteinen alue)

Todennäköisyys, että tuotteessa on jotain vialla:
P(A tai B)=0,1+0,2-0,05=0,25

Todennäköisyys, että tuotteessa on pelkästään käsittelyvika:

P(B, mutta ei A)=0,2-0,05=0,15

Todennäköisyys, että tuotteessa on täsmälleen yksi vika:

P(A tai B, muttei molemmat)=0,1+0,2-0,05-0,05=0,2

Kertolaskusääntö

Todennäköisyys, että sattuu tapahtuma A ja tapahtuma B:

P(A ja B) = P(A) · P(B|A)

Jos tapahtumat A ja B ovat toisistaan riippumattomia, niin

P(A ja B) = P(A) · P(B)

P(B|A) tarkoittaa B:n todennäköisyyttä sillä ehdolla, että A tapahtuu.

Esimerkki. Millä todennäköisyydellä kolmessa rahan heitossa saadaan kolme klaavaa (1. heitossa klaava ja 2. heitossa klaava ja 3. heitossa klaava)?

Rahanheitot ovat toisistaan riippumattomia, joten todennäköisyys on 0,5•0,5•0,5=0,125

Esimerkki. Erään liiketoimen onnistuminen edellyttää tapahtumien A ja B sattumista. Tapahtuman A todennäköisyys on 75 % ja tapahtuman B 70 %. Mikä on liiketoimen onnistumisen todennäköisyys?

Jos tapahtumat A ja B ovat toisistaan riippumattomia, niin onnistumisen todennäköisyys on 0,75•0,70=0,525=52,5 %.

Esimerkki. Suuressa tuote-erässä on 2 % virheellisiä. Millä todennäköisyydellä täsmälleen yksi satunnaisesti valituista kolmesta tuotteesta on virheellinen? Kyseessä on suuri tuote-erä, joten voimme olettaa että tuotteen virheellisyyden todennäköisyys ei riipu edellisen tuotteen virheellisyydestä.

- Todennäköisyys sille, että kolmesta tuotteesta ainoastaan ensimmäinen tuote on virheellinen, on 0,02·0,98·0,98
- Todennäköisyys sille, että ainoastaan toinen tuote on virheellinen, on 0,98·0,02·0,98
- Todennäköisyys sille, että ainoastaan kolmas tuote on virheellinen, on 0,98·0,98·0,02

Yhteenlaskusääntöä noudattaen todennäköisyys sille, että täsmälleen yksi on virheellinen, saadaan laskemalla:

0,02·0,98·0,98+0,98·0,02·0,98+0,98·0,98·0,02=3·0,02·0,98·0,98≈ 0,058=5,8 %

Esimerkki. 50 arvan joukossa on 10 voittoarpaa. Millä todennäköisyydellä täsmälleen yksi kolmesta ostetusta arvasta voittaa? Koska arpamäärä on suhteellisen pieni, niin arpojen voiton todennäköisyydet eivät ole toisistaan riippumattomia: voittoarvan todennäköisyys riippuu aiemmista arvoista. Pitää myös huomioida, että voittoarpa voi olla mikä tahansa kolmesta arvasta.

- Todennäköisyys sille, että 1. arpa voittaa ja muut kaksi arpaa eivät voita voidaan laskea kertolaskusäännöllä $10/50 \cdot 40/49 \cdot 39/48$.
- Todennäköisyys sille, että 2. arpa voittaa ja muut kaksi arpaa eivät voita voidaan laskea kertolaskusäännöllä $40/50 \cdot 10/49 \cdot 39/48$.
- Todennäköisyys sille, että 3. arpa voittaa ja muut kaksi arpaa eivät voita voidaan laskea kertolaskusäännöllä $40/50 \cdot 39/49 \cdot 10/48$.

Todennäköisyys sille että täsmälleen yksi arpa kolmesta voittaa saadaan edellä olevista yhteenlaskusäännöllä:

$10/50 \cdot 40/49 \cdot 39/48 + 40/50 \cdot 10/49 \cdot 39/48 + 40/50 \cdot 39/49 \cdot 10/48 \approx 0{,}398 = 39{,}8$ %

Esimerkki. Oletetaan, että lentokoneen moottorissa on moottorin toiminnan kannalta tärkeitä osia 1000 kpl. Millä todennäköisyydellä kone selviää lennosta, jos yksittäisen osan toimintavarmuudeksi lennon aikana oletetaan 99,9 %. Kertolaskusääntöä käyttäen:

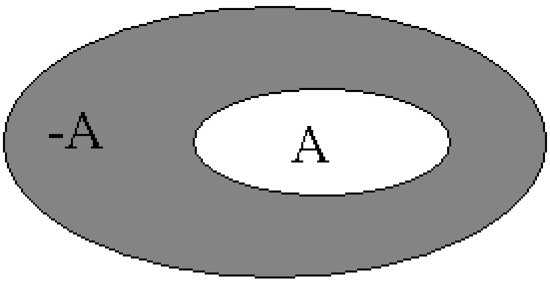
$0{,}999^{1000} \approx 0{,}368 = 36{,}8$ %

Komplementtisääntö

Joissain tapauksissa komplementtitapahtuman –A todennäköisyys on helpompi laskea kuin tapahtuman A todennäköisyys. Tällöin voidaan soveltaa komplementtisääntöä:

P(A) = 1 – P(-A)

Sääntö perustuu siihen tosiasiaan, että kaikkien alkeistapahtumien joukon todennäköisyys on 1. Vähentämällä kaikkien alkeistapahtumien joukon todennäköisyydestä A:n ulkopuolelle jäävien alkeistapahtumien todennäköisyys, jäljelle jää A:n todennäköisyys.



Esimerkki. Oleta ydinvoimalan toimintavarmuudeksi 99,99 % (todennäköisyys, ettei voimalan käyttöiän aikana satu vakavaa onnettomuutta). Mikä on todennäköisyys, että vähintään yhdessä voimalassa sattuu vakava onnettomuus, jos maapallolla on 500 ydinvoimalaa?

Kysyttyä todennäköisyyttä on vaikea laskea suoraan, koska onnettomuus voi sattua missä tahansa voimalassa ja onnettomuuksia voi sattua useammassakin ydinvoimalassa. Sen sijaan kannattaa käyttää hyväksi komplementtitapahtumaa: yhdessäkään voimalassa ei satu onnettomuutta.

$1-0{,}9999^{500} \approx 0{,}049 = 4{,}9\ \%$

Lasketaan vertailun vuoksi todennäköisyys, jos toimintavarmuudeksi oletetaankin 99,9 %

$1-0{,}999^{500} \approx 0{,}394 = 39{,}4\ \%$

Esimerkki. Arpajaisissa joka kolmas arpa voittaa. Ostetaan kolme arpaa. Millä todennäköisyydellä vähintään yksi arpa voittaa?

Jos A = vähintään yksi arpa voittaa, niin sen komplementti on

-A = yksikään arpa ei voita, jonka todennäköisyys on vaivatta laskettavissa.

$P(-A)=2/3\cdot 2/3\cdot 2/3$, joten $P(A)=1-2/3\cdot 2/3\cdot 2/3=19/27$

De Meren ongelmat: Vedonlyöntiä harrastava aatelismies de Mere pohdiskeli 1600-luvulla kannattaako lyödä vetoa seuraavien tapahtumien puolesta:

- vähintään yksi kuutonen neljässä nopanheitossa
- vähintään yksi kuutospari heitettäessä kahta noppaa 24 kertaa.

De Mere tuli pohdiskeluissaan päätelmään, jonka mukaan kummassakin tapauksessa voiton todennäköisyys olisi 2/3. Esimerkiksi ensimmäisessä tapauksessa de Mere laskeskeli 1/6+1/6+1/6+1/6=2/3. Aikansa vetoja lyötyään de Mere alkoi epäillä päättelyään. Mitä mieltä olet de Meren laskelmista?

BAYESIN KAAVA

Bayesin kaavan avulla korjataan aiemmin määriteltyä todennäköisyyttä käyttäen tarkasteltavaa ilmiötä koskevaa lisätietoa.

Kaupungissa on kaksi taksiliikennettä harjoittavaa yhtiötä: Sininen ja Vihreä. Toisen yhtiön autot ovat tummansinisiä ja toisen yhtiön autot tummanvihreitä. Sinisen takseja on 15 % ja Vihreän takseja 85 % kaikista kaupungin takseista. Eräänä pimeänä iltana taksiauto törmää jalankulkijaan ja jatkaa matkaansa törmäyksestä piittaamatta.

Määritellään tapahtumat B_1 ja B_2 seuraavasti:

- B_1=jalankulkijaan törmännyt taksi on sininen
- B_2=jalankulkijaan törmännyt taksi on vihreä.

Silminnäkijä sanoo nähneensä, että törmännyt taksi oli sininen. Jälkeenpäin testattiin samankaltaisissa ilta-olosuhteissa silminnäkijä-havaintojen luotettavuutta ja todettiin silminnäkijän tunnistavan auton värin oikein 80 % todennäköisyydellä.

Määritellään tapahtuma A=silminnäkijä sanoo nähneensä sinisen taksin.

Ilmiötä voidaan jäsentää taulukon avulla tarkastelemalla molempia mahdollisia tapahtumia (B_1 ja B_2) omilla riveillään:

Tapahtuma	**Alkuperäinen todennäköisyys**	**Silminnäkijän todennäköisyys**	**Yhdistetty todennäköisyys**	**Korjattu todennäköisyys**
Sininen törmäsi	$P(B_1)=0{,}15$	$P(A\|B_1)=0{,}80$	$P(A \text{ ja } B_1)=0{,}80\text{x}0{,}15=0{,}12$	$P(B_1\|A)=0{,}12/0{,}29=0{,}41$
Vihreä törmäsi	$P(B_2)=0{,}85$	$P(A\|B_2)=0{,}20$	$P(A \text{ ja } B_2)=0{,}20\text{x}0{,}85=0{,}17$	$P(B_2\|A)=0{,}17/0{,}29=0{,}59$
			$P(A)=0{,}12+0{,}17=0{,}29$	

Taulukossa on tarkasteltu molempia mahdollisuuksia (törmännyt taksi on sininen tai vihreä). Taulukon sarakkeeseen 'Yhdistetty todennäköisyys' on laskettu seuraavat tiedot:

- Jos törmännyt taksi on oikeasti sininen, niin silminnäkijä näkee taksin sinisenä 12 % todennäköisyydellä (laskenta kertolaskusääntöä käyttäen olettaen kaksi toisistaan riippuvaa tapahtumaa: törmännyt taksi on sininen ja silminnäkijä näkee taksin sinisenä)
- Jos törmännyt taksi on oikeasti vihreä, niin silminnäkijä näkee taksin sinisenä 17% todennäköisyydellä (laskenta kertolaskusääntöä käyttäen olettaen kaksi toisistaan riippuvaa tapahtumaa: törmännyt taksi on vihreä ja silminnäkijä näkee taksin sinisenä)
- Silminnäkijä näkee taksin sinisenä 29% todennäköisyydellä (laskenta yhteenlaskusääntöä käyttäen).

Taulukon viimeisessä sarakkeessa korjataan alkuperäistä todennäköisyyttä silminnäkijän antaman lisätiedon avulla:

- Törmännyt taksi on 41% todennäköisyydellä sininen olettaen, että silminnäkijä näki taksin sinisenä
- Törmännyt taksi on 59% todennäköisyydellä vihreä olettaen, että silminnäkijä näki taksin sinisenä.

Silminnäkijän antama lisätieto kasvatti siis todennäköisyyttä sille, että törmännyt taksi oli sininen arvosta 0,15 arvoon 0,41. Kuitenkin edelleen on todennäköisempää, että törmännyt taksi oli vihreä. Näppituntumalta moni arvioisi todennäköisyydet toisin ja antaisi silminnäkijälausunnolle enemmän painoa.

Elektroniikkalaitteita kokoava yritys tilaa erästä komponenttia kahdelta eri toimittajalta B_1 ja B_2. Toimittajalta B_1 tulee 65% komponenteista ja toimittajalta B_2 loput 35%. Aikaisemman kokemuksen mukaisesti toimittajan B_1 komponenteista 2 % on viallisia ja toimittajan B_2 komponenteista 5 %. Jos havaitaan virheellinen komponentti, niin korjatut todennäköisyydet toimittajan todennäköisyydestä ovat seuraavan taulukon mukaiset (A=tuote on viallinen):

Tapahtuma	**Alkuperäinen todennäköisyys**	**Viallisen todennäköisyys**	**Yhdistetty todennäköisyys**	**Korjattu todennäköisyys**
Toimittaja B_1	$P(B_1)=0{,}65$	$P(A\|B_1)=0{,}02$	$P(A \text{ ja } B_1)=0{,}02x0{,}65=0{,}0130$	$P(B_1\|A)=0{,}0130/0{,}0305=0{,}4262$
Toimittaja B_2	$P(B_2)=0{,}35$	$P(A\|B_2)=0{,}05$	$P(A \text{ ja } B_2)=0{,}05x0{,}35=0{,}0175$	$P(B_2\|A)=0{,}0175/0{,}0305=0{,}5738$
			$P(A)=0{,}0130+0{,}0175=0{,}0305$	

Bayesin kaava

Korjattu todennäköisyys voidaan laskea suoraan nk. Bayesin kaavaa käyttäen:

$$P(B\,|\,A)=\frac{P(B)\times P(A\,|\,B)}{P(A)}$$

Edellä olleessa esimerkissä korjattu todennäköisyys toimittajan B1 kohdalla olisi voitu laskea suoraan Bayesin kaavalla:

$$P(B_1\,|\,A)=\frac{0{,}65\times 0{,}02}{0{,}0305}=0{,}4262$$

Öljy-yhtiö suunnittelee porauksia alueella, jossa arvioidaan olevan korkealaatuista öljyä 50 % todennäköisyydellä (tapahtuma B_1), laadultaan keskinkertaista öljyä 20 % todennäköisyydellä (tapahtuma B_2) ja ei öljyä lainkaan 30 % todennäköisyydellä (tapahtuma B_3). Maaperästä porataan näytteitä. Näytteet ovat sellaisia, että aikaisemmissa tapauksissa vastaavanlaisia näytteitä on löydetty

- korkealaatuista öljyä sisältäviltä alueilta 20 % todennäköisyydellä
- keskilaatuista öljyä sisältäviltä alueilta 80 % todennäköisyydellä
- ei lainkaan öljyä sisältäviltä alueilta 20 % todennäköisyydellä.

Öljyn löytymiselle voidaan nyt laskea seuraavan taulukon mukaiset korjatut todennäköisyydet:

	Alkuperäinen todennäköisyys	**Porattu näyte**	**Yhdistetty todennäköisyys**	**Korjattu todennäköisyys**
Korkealaatuinen öljy	$P(B_1)=0{,}50$	$P(A\|B_1)=0{,}20$	$P(A \text{ ja } B_1)=0{,}10$	$P(B_1\|A)=0{,}10/0{,}32=0{,}3125$
Keskilaatuinen öljy	$P(B_2)=0{,}20$	$P(A\|B_2)=0{,}80$	$P(A \text{ ja } B_2)=0{,}16$	$P(B_2\|A)=0{,}16/0{,}32=0{,}5$
Ei öljyä	$P(B_3)=0{,}30$	$P(A\|B_3)=0{,}20$	$P(A \text{ ja } B_3)=0{,}06$	$P(B_3\|A)=0{,}06/0{,}32=0{,}1875$
			$P(A)=0{,}10+0{,}16+0{,}06=0{,}32$	

Näytteiden perusteella korkealaatuisen öljyn todennäköisyys putosi arvosta 0,50 arvoon 0,3125.

TODENNÄKÖISYYSJAKAUMAT

http://myy.haaga-helia.fi/~taaak/m/jakau.xlsx

Kahden kolikon heitossa tuloksena voi olla 0, 1 tai 2 klaavaa. Jos kahta kolikkoa heitetään 100 kertaa ja kirjataan muistiin klaavojen lukumäärä kullakin heittokerralla, niin tuloksena voi olla esimerkiksi:

klaavoja	% heitoista
0	29 %
1	53 %
2	18 %

Yllä olevaa jakaumaa voidaan kutsua empiiriseksi jakaumaksi.

Mahdollisia tuloksia kahden kolikon heitossa ovat

- 'kruunu-kruunu'
- 'kruunu-klaava'
- 'klaava-kruunu'
- 'klaava-klaava'.

Koska jokainen tulosvaihtoehto on yhtä todennäköinen, niin voimme päätellä klaavojen määrän teoreettiseksi jakaumaksi:

klaavoja	todennäköisyys
0	25%
1	50%
2	25%

Kyseistä jakaumaa voidaan kutsua todennäköisyysjakaumaksi.

Jos kahta kolikkoa heitetään kerta toisensa jälkeen, niin empiirinen jakauma alkaa vääjäämättä lähestyä todennäköisyysjakaumaa. Jos todennäköisyysjakauma ei ole pääteltävissä tai laskettavissa kuten edellä, niin todennäköisyysjakauman approksimaationa voidaan käyttää empiiristä jakaumaa tai todennäköisyysjakauman todennäköisyydet voidaan määrittää asiantuntija-arvioiden pohjalta (subjektiivisina todennäköisyyksinä).

Pitkävedossa voi olla yhtenä kohteena ottelu HIFK - Jokerit. Pitkävedossa veikataan ottelun lopputulosta jaotuksella kotivoitto, tasapeli tai vierasvoitto. Kertoimien määrittämiseksi Veikkaus Oy määrittänee todennäköisyysjakauman lopputulokselle (esimerkiksi kotivoitto 35 %, tasapeli 25 %, vierasvoitto 40 %). Todennäköisyydet pohjautuvat yhtäältä aiemmin pelattuihin peleihin ja toisaalta asiantuntija-arvioihin.

Todennäköisyysjakaumien yhteydessä käytetään käsitteitä:

- **satunnaismuuttuja**: muuttuja, jonka arvoihin liittyviä todennäköisyyksiä jakauma esittää (esimerkiksi klaavojen lukumäärä)
- **satunnaisilmiö**: tarkastelun kohteena oleva tapahtuma tai ilmiö (esimerkiksi kahden kolikon heitto)
- **diskreetti todennäköisyysjakauma**: satunnaismuuttujan mahdolliset arvot ja niiden todennäköisyydet.

Kertymätodennäköisyys

Kertymätodennäköisyys tarkoittaa todennäköisyyttä saada korkeintaan tietty satunnaismuuttujan arvo.

Kahden nopan heitossa (satunnaisilmiö) silmälukujen summa (satunnaismuuttuja) voi saada arvokseen 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 tai 12. Silmälukujen summan todennäköisyysjakauma selviää seuraavan ruudukon avulla:

	1	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Ruudukon yläreunassa on ensimmäisen nopan silmäluvut ja vasemmassa reunassa toisen nopan silmäluvut. Ruudukon soluihin on laskettu eri tulosvaihtoehtoihin liittyvät silmälukujen summat. Silmälukujen summa 2 voi ruudukon mukaan sattua vain yhdellä tavalla, siis todennäköisyys on 1/36. Silmälukujen summa 3 voi puolestaan sattua kahdella eri tavalla, siis todennäköisyys on 2/36 jne. Näin jatkaen voimme muodostaa todennäköisyysjakauman:

Silmälukujen summa	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Todennäköisyys	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36
Kertymätodennäköisyys	1/36	3/36	6/36	10/36	15/36	21/36	26/36	30/36	33/36	35/36	36/36

Kertymätodennäköisyys voidaan laskea satunnaismuuttujan arvon ja kaikkien sitä edeltävien satunnaismuuttujan arvojen todennäköisyyksien summana. Kertymätodennäköisyydet ovat keskeisessä asemassa todennäköisyysjakaumia käytettäessä. Eri jakaumien kertymätodennäköisyyksiä löytyy tilastotieteen kirjoista taulukoituna ja niitä voidaan laskea taulukkolaskennan funktioilla. Tämän vuoksi onkin syytä harjoitella erilaisten tapahtumien todennäköisyyksien laskemista kertymätodennäköisyyksien avulla.

Edellä kuvatussa kahden nopan silmälukujen jakaumassa:

- $P(X<7) = 15/36 = 5/12$ (todennäköisyys saada vähemmän kuin jotain selviää suoraan kertymätodennäköisyyden avulla)
- $P(X>9) = 1 – 30/36 = 6/36 = 1/6$ (todennäköisyys saada enemmän kuin jotain selviää kertymätodennäköisyyden komplementtina)
- $P(4<X<9) = 26/36 – 6/36 = 20/36 = 5/9$ (todennäköisyys saada jotain joltain väliltä selviää kertymätodennäköisyyksien erotuksena)

Odotusarvo

Empiiriselle jakaumalle voidaan laskea keskiarvo. Palataan aiempaan esimerkkiin, jossa heitettiin kahta kolikkoa 100 kertaa ja klaavojen lukumäärän jakaumaksi saatiin:

klaavoja	% heitoista
0	29 %
1	53 %
2	18 %

Koska heittoja oli 100 kappaletta, niin klaavojen kappalemäärät ovat samoja kuin prosenttiosuudet. Näin ollen keskiarvoksi saadaan

$$\frac{29\cdot 0+53\cdot 1+18\cdot 2}{100}=0{,}89$$

Voimme myös määrittää keskiarvon eli odotusarvon todennäköisyysjakaumalle. Jos todennäköisyysjakauma tunnetaan, niin **odotusarvo on satunnaismuuttujan arvojen todennäköisyyksillä painotettu summa.**

Aiemmin totesimme kahden kolikon heitossa klaavojen lukumäärän todennäköisyysjakaumaksi:

klaavoja	todennäköisyys
0	25 %
1	50 %
2	25 %

Klaavojen lukumäärän odotusarvo on klaavojen lukumäärien todennäköisyyksillä painotettu summa:

$0{,}25\cdot 0+0{,}5\cdot 1+0{,}25\cdot 2=1$

Odotusarvot ovat keskeisiä monissa sovelluksissa. Esimerkiksi vakuutusyhtiöt ovat kiinnostuneita tulevan vuoden vakuutuskorvausten odotusarvoista eri vakuutustyyppien kohdalla, sijoittajat ovat kiinnostuneita sijoitusten tuottojen odotusarvoista ja kauppiaat ovat kiinnostuneita tuotteiden kysynnän odotusarvoista.

Esimerkki. Suunnitelmissa olevan investoinnin osalta arvioidaan:

- investoinnin kustannukset ovat seuraavan 10 vuoden aikana 100 000 euroa vuodessa
- korkeasuhdanteessa saadaan tuottoja noin 180 000 euroa vuodessa
- matalasuhdanteessa saadaan tuottoja 110 000 euroa vuodessa
- korkeasuhdanteen todennäköisyys on tilastojen mukaan 0,40 ja matalasuhdanteen 0,60

Tietojen pohjalta voimme laskea investoinnin vuosituoton odotusarvon:

$0{,}40\cdot 180\,000+0{,}60\cdot 110\,000=138\,000$

Tuoton odotusarvon avulla voidaan sitten arvioida onko investointi riittävän kannattava.

Esimerkki. Arpajaisissa on 1000 arpaa. Voittoarpoja on 31:

- 500 euron voittoja on yksi
- 300 euron voittoja on 10
- 100 euron voittoja on 20

Määritä arvan hinta siten, että voiton odotusarvo on 55 % arvan hinnasta.

Lasketaan voiton odotusarvo:

$$\frac{1}{1000}\cdot 500+\frac{10}{1000}\cdot 300+\frac{20}{1000}\cdot 100=5{,}50$$

Jotta voiton odotusarvo 5,50 euroa olisi 55 % arvan hinnasta, täytyy arvan hinnaksi asettaa 10 euroa.

Esimerkki. Maahantuojalla on tilasto (perustuen aiemmin tehtyihin korjauksiin) automalliin takuuaikana tehtävistä korjauksista:

- 50% autoista ei tarvitse tehdä takuun piiriin kuuluvia korjauksia
- 20% autoista joudutaan tekemään keskimäärin 150 euroa maksavat korjaukset
- 25% keskimäärin 400 euroa maksavat korjaukset
- loppuihin 600 euroa maksavat korjaukset.

Tietojen pohjalta voidaan määrittää, kuinka paljon takuun pitäisi lisätä uuden auton hintaa. Riittää, kun laskemme korjauksiin kuluvan euromäärän odotusarvon:

$0{,}20\cdot 150+0{,}25\cdot 400+0{,}05\cdot 600=160$

Esimerkki. Vakuutusyhtiöissä hyödynnetään laajasti erilaisiin ilmiöihin liittyviä odotusarvoja vakuutuksia hinnoiteltaessa. Tarkastellaan esimerkkinä yksinkertaista ulkoilmatapahtuman järjestäjälle myönnettävää sadevakuutusta. Ulkoilmatapahtuman järjestäjä ottaa tällaisen vakuutuksen lieventääkseen mahdollisen sateen aiheuttamia tappioita. Kiinnekohta, jonka perusteella vakuutus voidaan hinnoitella, on kyseisen ajankohdan sateen todennäköisyysjakauma.

	A	B	C
1	**sademäärä (mm)**	**todennäköisyys**	**korvaus**
2	0-2	53 %	-
3	3-5	30 %	500
4	6-10	15 %	1000
5	11-	2 %	3000
6			
7	**Korvauksen odotusarvo**	**360**	

Jos laskelmat tehdään taulukkolaskentaa käyttäen, niin yllä olevan taulukon tapauksessa odotusarvo voidaan laskea kaavalla **=SUMPRODUCT(B3:B5;C3:C5)** (TULOJEN.SUMMA). Taulukkolaskentamallissa on helppo kokeilla erilaisia korvaussummia eri sademäärille ja seurata odotusarvon muutoksia.

Vakuutusyhtiö tietenkin lisää odotusarvoon oman katteensa.

Binomijakauma

Oletetaan onnenpyörä, jota pyöräyttämällä voittaa 15 % todennäköisyydellä. Useampaan onnenpyörän pyöräytykseen voidaan liittää todennäköisyysjakauma käyttäen voittojen lukumäärää satunnaismuuttujana. Esimerkiksi viidelle pyöräytykselle saadaan jakaumaksi:

voittojen määrä	todennäköisyys
0	44,3705 %
1	39,1505 %
2	13,8178 %
3	2,4384 %
4	0,2152 %
5	0,0076 %

Jakauman todennäköisyyksien laskeminen perustuu binomijakaumaan. Binomijakaumaa voidaan soveltaa, jos satunnaisilmiö ja satunnaismuuttuja toteuttavat seuraavat ehdot:

- Satunnaisilmiötä toistetaan useita kertoja (toistojen määrää merkitään n)
- Satunnaismuuttujan arvot voidaan jakaa täsmälleen kahteen luokkaan, joiden todennäköisyydet tiedetään. Jos toisen todennäköisyyttä merkitään p, niin toisen todennäköisyys on 1-p.
- Satunnaismuuttujan arvojen todennäköisyydet pysyvät vakioina toistokerrasta toiseen.

Edellä tarkasteltu onnenpyörä toteuttaa mainitut ehdot:

- Satunnaisilmiötä toistetaan 5 kertaa
- Satunnaismuuttujalla on kaksi arvoa: voitto, ei voittoa
- Voiton todennäköisyys on vakio p = 15 %.

Binomijakauma-kaavalla voidaan laskea todennäköisyys sille, että satunnaisilmiötä n kertaa toistettaessa saadaan k kappaletta onnistumisia (onnistumisella tarkoittaa jompaakumpaa satunnaismuuttujan luokista). Binomijakauma-kaava on:

$$\frac{n!}{k!\cdot(n-k)!}\cdot p^k \cdot (1-p)^{n-k}$$

Esimerkiksi todennäköisyys saada täsmälleen 2 voittoa onnenpyörän viidessä pyörityksessä lasketaan seuraavasti (n=5, k=2, p=0,15):

$$\frac{5!}{2!\cdot(5-2)!}\cdot 0{,}15^2 \cdot 0{,}85^3 \approx 0{,}138178 = 13{,}8178\,\%$$

Kaava voi näyttää mutkikkaalta, mutta sille löytyy selitys. Edellä kaavan alussa lasketaan kombinaatio-kaavalla kuinka monella tavalla 2 voittoa voi esiintyä 5 pyörityksen joukossa. Kaavan jälkimmäisessä osassa lasketaan yksittäisen kahden voiton kombinaation todennäköisyys tuloperiaatteella (2 voittoa, 3 ei-voittoa).

Muita esimerkkejä binomijakautuneista satunnaismuuttujista:

- klaavojen määrä heitettäessä kolikkoa 10 kertaa
- kuutosten määrä heitettäessä kahta noppaa 24 kertaa
- voittojen määrä ostettaessa 6 arpaa arpajaisista, joissa joka kolmas arpa voittaa
- punaisten lukumäärä 15 ruletin pyörityksessä

- viallisten lukumäärä viiden tuotteen erässä
- ydinvoiman kannattajien määrä 1000 henkilön otoksessa
- ostavien asiakkaiden määrä sisään saapuneista 100 asiakkaasta.

Esimerkki. Joka kolmas arpa voittaa. Katsotaanpa miten käy, kun ostetaan kuusi arpaa. Voittojen lukumäärä noudattaa binomijakaumaa:

- arvan osto toistetaan 6 kertaa (n=6)
- arpa voi voittaa tai sitten ei, voiton todennäköisyys p=1/3
- Voiton todennäköisyys pysyy vakiona (oletetaan, että arpaerä on niin iso, ettei voittoarvon nostaminen olennaisesti muuta voiton todennäköisyyttä seuraavilla arvoilla).

	A	B	C	D
1	**Arvat**			
2				
3	Arpoja ostetaan (n)	6		
4	Voiton todennäköisyys (p)	33,3 %		
5				
6	**Todennäköisyysjakauma**	Voittoja	Todennäköisyys	Kertymä todennäköisyys
7		0	8,8 %	8,8 %
8		1	26,3 %	35,1 %
9		2	32,9 %	68,0 %
10		3	21,9 %	90,0 %
11		4	8,2 %	98,2 %
12		5	1,6 %	99,9 %
13		6	0,1 %	100,0 %

Edellä olevassa Excel-taulukossa solun C7 todennäköisyys voidaan laskea funktiolla **=BINOMDIST(B7;$B$3;$B$4;0)** (BINOMIJAKAUMA) ja solun D7 todennäköisyys funktiolla =BINOMDIST(B7;$B$3;$B$4;1). Viimeinen argumentti siis määrittää lasketaanko todennäköisyyttä vai kertymätodennäköisyyttä. Funktioita alaspäin kopioimalla saadaan muiden rivien todennäköisyydet.

Esimerkki Kyselytutkimuksessa käytettävään otokseen voi sattumalta tulla hyvinkin erilainen koostumus kuin kiinnostuksen kohteena olevassa perusjoukossa. Binomijakaumaa käyttäen voidaan helposti laskea kuinka todennäköistä on saada kiinnostuksen kohteena olevaa joukkoa vastaava otos. Oletetaan, että 50 % perusjoukosta vastustaa uutta ydinvoimalaa. Nyt voimme laskea esimerkiksi kuinka suuri on todennäköisyys, että 1000 henkilön satunnaisotokseen sattuu 47 % - 53 % ydinvoiman vastustajia?

Funktiolla =BINOMDIST(530;1000;50%;1) selviää, että todennäköisyys korkeintaan 53 prosentille on noin 97,3 %. Funktiolla =BINOMDIST(470;1000;50%;1) selviää, että todennäköisyys korkeintaan 47 prosentille on noin 3,1 %. Näiden erotuksena saadaan todennäköisyys 94,2 % sille että vastustajia sattuu otokseen 47 % - 53 % .

Esimerkki. Hotellit ja lentoyhtiöt ottavat yleisesti varauksia enemmän kuin paikkoja on tarjolla. Tämä perustuu kokemukseen siitä, että kaikki paikan varanneet eivät kuitenkaan saavu paikalle. No show -tapausten määrä vaihtelee, joten varmuudella ei voida laskea soveliasta ylimääräisten varausten lukumäärää. Todennäköisyysjakaumaa käyttäen voidaan kuitenkin laskea ylimääräisten varausten sovelias lukumäärä, jos esimerkiksi halutaan 95 % varmuus paikkojen riittävyydelle.

	A	B
1	**Lentovuoron tai hotellin ylivaraukset**	
2	Voit muuttaa kehystettyjä arvoja	
3		
4	Paikkoja	400
5		
6	No-show todennäköisyys	5 %
7	Show todennäköisyys	95 %
8		
9	Varauksia	414
10		
11	Todennäköisyys, että paikat riittävät	95,5 %
12	Todennäköisyys, että paikat eivät riitä	4,5 %

Edellä olevaa taulukkoa käyttäen voidaan laskea riski paikkojen riittämättömyydelle. Taulukon lähtötietoja (paikkojen määrä, no-show todennäköisyys) voi vaihdella. Taulukon solussa B11 on funktio =BINOMDIST(B4;B9;B7;1) ja solussa B12 on kaava =1-B11.

Poisson-jakauma

Kun binomijakauman toistojen määrää n kasvatetaan ja todennäköisyyttä p pienennetään, niin binomijakauman todennäköisyyden kaava lähestyy raja-arvoa:

$$e^{-\lambda}\frac{\lambda^k}{k!}$$

Yllä olevan kaavan määrittelemää Poisson-jakaumaa voidaan käyttää jos toistojen määrä n on suuri ja tarkasteltava tapahtuma on harvinainen (todennäköisyys p on pieni). Toistojen määrän ei itse asiassa tarvitse edes olla tiedossa. Ainoa kaavassa tarvittava tieto on odotusarvo λ, joka on tapahtumien keskimääräinen lukumäärä (kuinka monta kertaa tapahtuma keskimäärin on sattunut tarkasteltavana olevassa erässä, aikavälissä jne.). Kaavassa esiintyvä e on luonnollisen logaritmijärjestelmän kantaluku, jonka likiarvo on 2,718. Yleisimmin Poisson-jakaumaa sovelletaan tietyssä aikavälissä sattuvien tapahtumien lukumäärän tarkasteluun.

Esimerkki. Pankin konttoriin lounasaikana viiden minuutin aikana saapuvien asiakkaiden määrän voidaan olettaa olevan Poisson-jakautunut. Toistojen määrä on suuri (kaikki pankin asiakkaat) ja todennäköisyys, että tietty satunnaisesti valittu pankin asiakas saapuu juuri kyseisen viiden minuutin aikana, on pieni. Poisson-jakauman käyttämiseksi lounasaikana saapuvien asiakkaiden määrää pitää tarkkailla viiden minuutin ajanjaksoissa, josta saadaan laskettua keskimäärin viidessä minuutissa saapuvien asiakkaiden määrä.

Esimerkki. Tieliikenteessä vuosittain kuolevien lukumäärää voidaan mallintaa Poisson-jakaumalla. Toistojen määrä on suuri (kaikki tieliikenteessä liikkuvat) ja todennäköisyys, että tietty satunnaisesti valittu henkilö kuolee juuri kyseisen vuoden aikana tieliikenteessä, on pieni. Poisson-jakauman käyttämiseksi tarvitaan historiatietoja, joiden perusteella voidaan laskea keskimäärin vuodessa tieliikenteessä kuolleiden lukumäärä.

Esimerkki. Ensiapuaseman johtajaa saattaa kiinnostaa tietää kuinka suurta potilasmäärää ensiapuasemalle voidaan tietyssä työvuorossa lauantai-iltana odottaa, jotta henkilökunnan määrä voitaisiin mitoittaa sopivaksi. Aikaisempien viikkojen ja kuukausien tietoja tarkastelemalla selviää, että keskimäärin ensiapuasemalle on saapunut lauantai-iltaisin 3 potilasta tunnissa. Johtaja ajattelee, että juurihan viime viikolla varauduttiin kolmen potilaan käsittelyyn tunnissa, mutta tuloksena oli täysi kaaos ja joukko vihaisia odotushuoneessa odottavia potilaita. Tarkastellaanpa pelkän keskiarvon sijasta todennäköisyysjakaumaa: Tietyssä aikavälissä saapuvien asiakkaiden määrän voidaan olettaa noudattavan Poisson-jakaumaa.

	A	B	C	D	E
1	**Asiakkaiden saapuminen palvelupisteeseen**				
2	Voit muuttaa kehystettyä arvoa				
3					
4	Keskiarvo	3			
5					
6	**Todennäköisyysjakauma**	asiakkaita	todennäköisyys	kertymä todennäköisyys	todennäköisyys isommalle asiakasmäärälle
7		0	5,0 %	5,0 %	95,0 %
8		1	14,9 %	19,9 %	80,1 %
9		2	22,4 %	42,3 %	57,7 %
10		3	22,4 %	64,7 %	35,3 %
11		4	16,8 %	81,5 %	18,5 %
12		5	10,1 %	91,6 %	8,4 %
13		6	5,0 %	96,6 %	3,4 %
14		7	2,2 %	98,8 %	1,2 %

Taulukon solussa C7 on funktio **=POISSON(B7;$B$4;0)** ja solussa D7 funktio =POISSON(B7;$B$4;1). Huomaa, että funktion viimeinen argumentti ilmaisee sen lasketaanko yksittäiseen lukumäärään liittyvä todennäköisyys (0) vai kertymätodennäköisyys (1). Funktioita alaspäin kopioimalla saadaan muiden rivien todennäköisyydet. Jos varaudutaan 6 potilaan vastaanottoon, niin taulukon mukaan riski isommalle potilasmäärälle on 3,4 %.

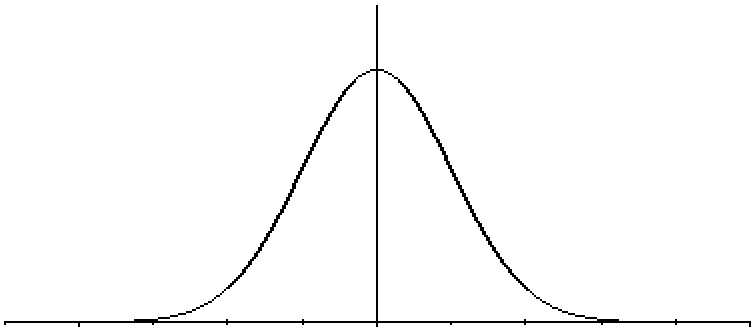
Esimerkki. Kodinkoneliike haluaa arvioida kuinka monta tietyn mallista TV:tä kannattaa pitää varastossa, kun tiedetään keskimääräiseksi viikkomyynniksi 5 kpl. Edellisen esimerkin taulukosta saadaan pienin muutoksin tähän esimerkkiin sopiva taulukko. Keskimääräisen potilasmäärän 3 tilalle kirjoitetaan TV:n keskimääräinen myyntimäärä 5. Taulukosta voidaan tämän jälkeen etsiä esimerkiksi lukumäärä, joka takaa riittävyyden 90 % varmuudella.

Normaalijakauma

Tähtitiede kehittyi ripein askelin 1500-luvulta lähtien. Tutkimusta haittasivat kuitenkin epätarkat mittaustulokset (esimerkiksi taivaankappaleiden etäisyyksiä, sijaintia ja liikkeitä mitattaessa). Epätarkkuus johtui mm. mittauslaitteiden kehittymättömyydestä ja inhimillisistä tekijöistä. Jo Galileo Galilei tuli tulokseen, että mittausvirheet noudattavat symmetristä jakaumaa ja että pieniä virheitä esiintyy useammin kuin suuria. Gauss niminen saksalainen matemaatikko ja tähtitieteilijä esitti vuonna 1809 lausekkeen jakaumalle, jota mittausvirheet noudattivat. Nykyään kyseistä jakaumaa kutsutaan normaalijakaumaksi.

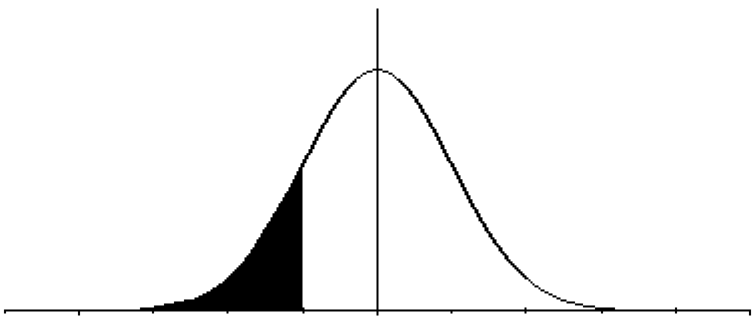
Normaalijakauman lauseke määrittelee funktion, jolla on seuraavia ominaisuuksia:

- se on satunnaismuuttujan arvon funktio
- funktion arvojen laskemiseksi täytyy tietää jakauman keskiarvo (odotusarvo) ja keskihajonta
- funktion kuvaaja on symmetrinen ns. Gaussin kellokäyrä, jonka huippu on keskiarvon kohdalla
- funktion alle jäävä pinta-ala voidaan samaistaa todennäköisyyteen.



On tärkeää huomata, että normaalijakauman yhteydessä ei ole mielekästä puhua yksittäisen satunnaismuuttujan arvon todennäköisyydestä. Normaalijakaumaa noudattava satunnaismuuttujahan on jatkuva eli voi saada mitä tahansa arvoja tietyltä väliltä. Tällöin yksittäiseen arvoon liittyvä todennäköisyys on ainakin teoreettisesti ajatellen 0. Yksittäisten arvojen todennäköisyyksien sijasta on mielekästä puhua kertymätodennäköisyyksistä ja kertymätodennäköisyyksien avulla voidaan edelleen laskea erilaisten välien todennäköisyyksiä.

Käytännössä normaalijakaumaa hyödynnetään kertymätodennäköisyyksien avulla. Satunnaismuuttujan arvoon x liittyvä kertymätodennäköisyys on kohdan x vasemmalle puolella oleva käyrän alapuolelle jäävä pinta-ala.



Eri satunnaismuuttujan arvoihin liittyviä kertymätodennäköisyyksiä voidaan laskea taulukkolaskentafunktiolla **=NORMDIST(x;odotusarvo;keskihajonta;1)** (NORM.JAKAUMA). Kertymätodennäköisyyksiä on myös taulukoitu tilastotieteen kirjoihin.

Esimerkki. Älykkyysosamäärää voidaan mitata tarkoitusta varten laaditulla testillä. Testin pisteytys on skaalattu sellaiseksi, että amerikkalaisten älykkyysosamäärä noudattaa normaalijakaumaa N(100,16) (merkintä tarkoittaa normaalijakaumaa, jonka odotusarvo on 100 ja keskihajonta 16).

Älykkyysosamäärään 80 liittyvä kertymätodennäköisyys on noin 10,6 %. Tämä selviää taulukkolaskentafunktiolla =NORMDIST(80;100;16;1). Siis 10,6 % amerikkalaisista on älykkyysosamäärältään alle 80. Vastaavasti 89,4 % amerikkalaisista on älykkyysosamäärältään yli 80.

Todennäköisyys, että henkilön älykkyysosamäärä on täsmälleen 80, on ainakin teoreettisesti ajatellen 0. Käytännössä näin ei ehkä ole, koska älykkyysosamäärätestin pisteytys ei ole mielivaltaisen tarkkaa. Jos esimerkiksi testi pisteytetään pisteen tarkkuudella, niin todennäköisyys älykkyysosamäärälle 80 on varmasti selvästi nollasta poikkeava. Normaalijakaumaa käytettäessä kuitenkin oletetaan yksittäiseen satunnaismuuttujan arvoon liittyväksi todennäköisyydeksi 0.

Todennäköisyys, että henkilön älykkyysosamäärä on välillä 80 - 110 saadaan selville vähentämällä arvoon 110 liittyvästä kertymätodennäköisyydestä (noin 73,4 %) arvoon 80 liittyvä kertymätodennäköisyys (noin 10,6 %). Tulokseksi saadaan noin 62,8 %.

Entäpä jos halutaan tietää minkä älykkyysosamäärän ylittää älykkäin kymmenesosa amerikkalaisista? Kysytään siis satunnaismuuttujan arvoa, kun todennäköisyys tunnetaan. Tällaisten ongelmien ratkaisuun voidaan käyttää taulukkolaskentafunktiota
=NORMINV(kertymätodennäköisyys,odotusarvo,keskihajonta)
(NORM.JAKAUMA.KÄÄNT). Funktio =NORMINV(90%;100;16) antaa vastauksen 120,5.

Esimerkki. Oletetaan, että aikaisemmista myyntitilastoista voidaan todeta, että päivittäistavaran kysynnällä lauantaisin on ollut keskiarvo 120 ja keskihajonta 15. Kauppias voi käyttää normaalijakaumaa määrittäessään seuraavan lauantain tilausmäärää kyseisen päivittäistavaran kohdalla:

	A	B
1	Keskiarvo	120
2	Keskihajonta	15
3	Varmuus	95 %
4		
5	Tilausmäärä	145

Ylläkuvatussa laskentamallissa voidaan käyttää solussa B5 funktiota =NORMINV(B3;B1;B2), joka palauttaa 145. Jos kauppias haluaa 95 % varmuuden tavaran riittävyydelle, niin hänen kannattaa tilata 145 kappaletta. Kauppias voi tietenkin käyttää varmuusprosenttina muutakin kuin 95 %. Käytettyyn taulukkolaskentamalliin on helppo muuttaa lähtötietoja ja laskea tilausmääriä myös muille päivittäistavaroille.

Esimerkki. Paperikone voidaan säätää valmistamaan eri painoisia paperilaatuja (80 grammaa/neliömetri, 85 grammaa/neliömetri jne.). Teollisesti valmistetun tuotteen ominaisuudet kuitenkin vaihtelevat monien satunnaisten tekijöiden takia. Vaikka kone on säädetty valmistamaan 80 g paperia, niin paperi ei ole kauttaaltaan 80 g. Oletetaan, että paperin tilaaja vaatii, että 90 % paperista täytyy olla yli 80 g. Minkälaista paperia paperikone pitäisi säätää valmistamaan (paperin neliömetripainon odotusarvo on siis säädettävissä). Oletetaan, että valmistusprosessiin ja kyseiseen paperikoneeseen liittyvä keskihajonta on 2,5 grammaa/neliömetri. Tehtävä voidaan ratkaista taulukkolaskentamallilla:

	A	B
1	Keskiarvo	83,2100
2	Keskihajonta	2,5000
3		
4	Satunnaismuuttujan arvo	80,0000
5	Todennäköisyys saada vähemmän	9,9571 %
6	Todennäköisyys saada enemmän	90,0429 %

Solussa B5 on funktio =NORMDIST(B4;B1;B2;1) ja solussa B6 kaava =1-B5. Solun B5 arvoja vaihtelemalla voidaan etsiä arvoa, joka antaa 90 % todennäköisyyden saada yli 80 g paperia. Yllä kokeilulla on etsitty sopiva keskiarvo kahden desimaalin tarkkuudella. Jos käytännön sovelluksessa vaaditaan suurempaa tarkkuutta, niin voidaan jatkaa kokeilua kolmannen desimaalin kohdalla.

Esimerkki. Var eli Value at risk on yleisesti käytetty työkalu sijoituskohteiden riskin arviointiin. Jos tarkastellaan päiväkohtaisia Var-lukuja, niin esim. 5 % Var-luku on sellainen päivätuotto, jota heikomman päivätuoton todennäköisyys on 5 %. Jos sijoituskohteen päivätuottojen oletetaan noudattavan normaalijakaumaa, niin Var-arvo voidaan määrittää jakaumasta. Monilla sijoitusrahastoilla on kirjattu rahaston sääntöihin rajoituksia Var-luvuille (esimerkiksi rahaston 5 % Var-luku ei saa olla pienempi kuin -2 %). Tällaiset säännöt sitovat sijoitusrahaston hoitajaa. Jos jonain päivänä Var-luku ei ole sääntöjen mukainen, niin sijoitusrahaston kokoonpanoa joudutaan vaihtamaan vähäriskisempiin.

Oletetaan, että sijoitusrahaston päivätuotto noudattaa normaalijakaumaa N(0,06 %;2 %-yksikköä). Tällöin 5 % Var-luku saadaan taulukkolaskentafunktiolla =NORMINV(5%;0,06;2). Tulokseksi tulee noin -3,2 %. Sijoitusrahaston päivätappio voi siis 5 % todennäköisyydellä olla -3.2% tai enemmän. Jos rahaston sääntöjen mukaan Var-luku ei saa alittaa -2 %, niin rahaston hoitajan on ryhdyttävä toimenpiteisiin riskin pienentämiseksi.

Edellä olleista esimerkeistä nähdään normaalijakauman soveltuvan malliksi hyvin erilaisiin käytännön tilanteisiin. Yhteistä edellä tarkastelluille satunnaismuuttujille on, että niiden arvoihin vaikuttavat monet satunnaiset tekijät. Tämä pätee yleisemminkin: jos satunnaismuuttujan arvo määräytyy satunnaisesti lukuisten eri tekijöiden vaikutuksesta, niin muuttuja yleensä noudattaa likimain normaalijakaumaa.

Normitettu normaalijakauma

Normaalijakaumaa, jonka keskiarvo on 0 ja keskihajonta 1 kutsutaan normitetuksi normaalijakaumaksi. Jos normaalijakauman todennäköisyyksiä määritetään tilastotieteen kirjoihin taulukoitujen kertymätodennäköisyyksien avulla, niin normitetun jakauman tunteminen on tärkeää. Taulukoidut todennäköisyydet ovat nimenomaan normitetun jakauman todennäköisyyksiä. Eri keskiarvon ja keskihajonnan omaavien jakaumien välillä on yhteys: Jos kahdessa normaalijakaumassa ollaan yhtä monen keskihajonnan päässä keskiarvosta, niin kertymätodennäköisyydet ovat samat.

Esimerkki. Normitetussa jakaumassa N(0,1) todennäköisyys saada korkeintaan -1 on noin 15,9 %. Tämän perusteella jakaumassa N(120,20) todennäköisyys saada korkeintaan 100 on myös 15,9 %. Molemmissa jakaumissahan tarkastellaan jakauman kohtaa, joka on yhden keskihajonnan päässä jakauman keskiarvosta.

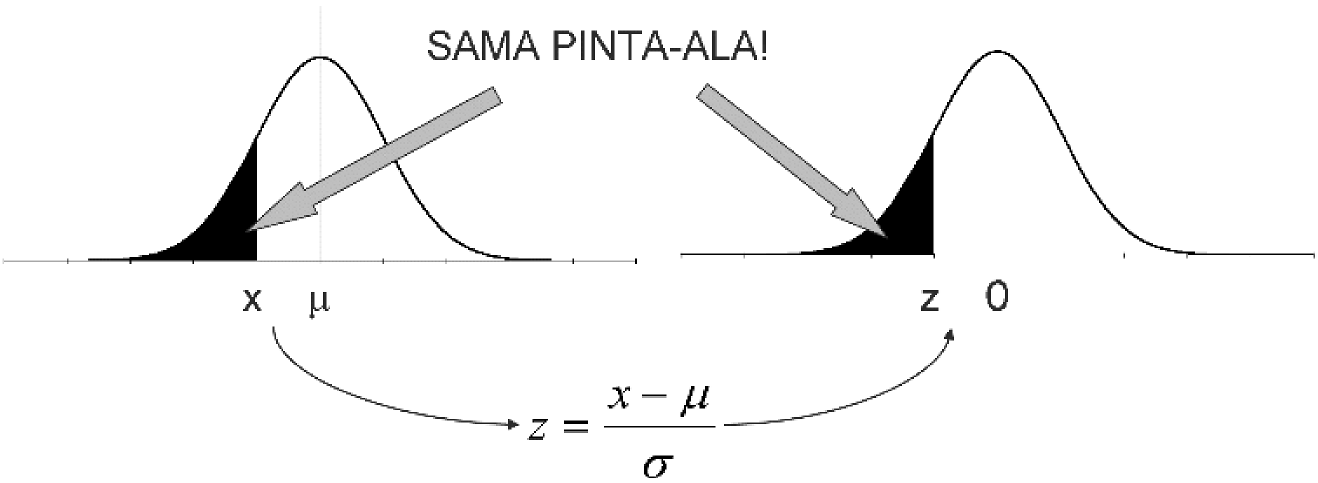
Joitain kaikille normaalijakaumille päteviä todennäköisyyksiä:

- Todennäköisyys saada jotain väliltä keskiarvo +/- keskihajonta on noin 68 %
- Todennäköisyys saada jotain väliltä keskiarvo +/- 2 keskihajontaa on noin 95 %
- Todennäköisyys saada jotain väliltä keskiarvo +/- 3,3 keskihajontaa on noin 99 %

Normitettu satunnaismuuttuja tarkoittaa satunnaismuuttujaa, jonka arvot on muunnettu normitettua jakaumaa vastaaviksi. Muunnoksessa on olennaista todennäköisyyksien säilyttäminen samoina. Edellä jo todettiin, että todennäköisyydet säilyvät kunhan pysytään yhtä monen keskihajonnan päästä keskiarvosta.

Esimerkki. Jakauman N(120,20) satunnaismuuttujan 100 normitettu arvo on -1, koska 100 on yhden keskihajonnan verran keskiarvon vasemmalla puolella.

Yleinen kaava satunnaismuuttujan normittamiselle on:



Kaavan yläosassa lasketaan vähennyslaskulla satunnaismuuttujan etäisyys keskiarvosta. Tulos jaetaan keskihajonnalla (käytetään perusjoukon keskihajontaa jos se tiedetään).

Satunnaismuuttujan normittamista käytetään monissa tilastollisissa analyyseissä, mutta myös haluttaessa saada erilaiset muuttujat vertailukelpoisiksi keskenään.

Eksponentiaalinen jakauma

Eksponentiaalisella jakaumalla on yhteys Poisson-jakaumaan. Jos tapahtumien lukumäärä tietyllä aikavälillä noudattaa Poisson-jakaumaa, niin tapahtumien välinen aika noudattaa eksponentiaalista jakaumaa. Tavallisia eksponentiaalisen jakauman sovelluskohteita ovat:

- palvelupisteeseen saapuvien asiakkaiden väliaika
- palvelun kestoaika
- komponentin kestoikä.

Lähtötietona tarvitaan tapahtumien välisen ajan keskiarvo μ. Eksponentiaalisen jakauman kertymätodennäköisyys kohdassa x saadaan kaavasta

$$1-e^{-x/\mu}$$

missä e on luonnollisen logaritmijärjestelmän kantaluku, jonka likiarvo on 2,718 ja μ on tapahtumien välisen ajan keskiarvo. Jos esimerkiksi asiakkaita saapuu keskimäärin 3 minuutin välein, niin seuraavan asiakkaan saapumiseen kuluu korkeintaan 2 minuuttia todennäköisyydellä

$$1 - e^{-2/3}$$

eli noin 0,4866 (48,66 %).

Excelissä eksponentiaalisen jakauman kertymätodennäköisyyksiä lasketaan funktiolla

=EXPONDIST(x;1/µ;1) (EKSPONENTIAALIJAKAUMA)

Äskeisen esimerkin kertymätodennäköisyys saadaan siis funktiolla

=EXPONDIST(2;1/3;1).

Excelissä ei ole erillistä funktiota satunnaismuuttujan arvon laskemiseksi, kun kertymätodennäköisyys tunnetaan. Satunnaismuuttujan arvo voidaan kuitenkin helposti ratkaista kertymäfunktion lausekkeesta, jolloin saadaan:

$$x = -\mu * \ln(1 - P)$$

missä P on kertymätodennäköisyys ja ln luonnollinen logaritmi.

SIMULOINTI

Monia satunnaisilmiöitä voidaan mallintaa todennäköisyysjakaumien avulla. Jo yksinkertaisissakin tapauksissa on usein vaikeaa laskea todennäköisyysjakaumien pohjalta päätöksentekoa ohjaavia ratkaisuja. Varteenotettava vaihtoehto on ilmiön jäljitteleminen eli simulointi.

Simulointi perustuu todennäköisyysjakaumista arvottaviin satunnaislukuihin. Tietokonetta käyttäen voidaan vaivatta simuloida samaa ilmiötä kerta toisensa jälkeen ja tilastoida eri simulointikertojen lopputulemat. Yhteenvetona voidaan laskea tilastollisia tunnuslukuja, jotka antavat tietoa tarkasteltavasta ilmiöstä.

Esimerkki. Tavaratalon kassoille tiettyyn aikaan päivästä saapuvien henkilöiden määrää voidaan onnistuneesti mallintaa todennäköisyysjakaumalla. Samoin kassalle saapuvan henkilön palvelemiseen kuluvaa aikaa voidaan mallintaa todennäköisyysjakaumalla. On kuitenkin vaikea laskea mitään analyyttistä ratkaisua sopivaksi kassojen lukumääräksi. Simuloinnin avulla voidaan kuitenkin luoda kuvitteellinen tilanne, jossa arvotaan todennäköisyysjakaumasta asiakkaita kassoille ja edelleen arvotaan kyseisille asiakkaille palveluajat. Simulointia voidaan jatkaa, kunnes saadaan riittävän tarkka kuva jonon muodostuksesta.

Seuraavassa esitellään esimerkkeinä joitain yksinkertaisia simulointimalleja, joihin voit tutustua lähemmin Excel-esimerkkien avulla.

Satunnaislukujen arvonta Excelissä

Excelissä on satunnaislukuja arpova funktio **RAND** (SATUNNAISLUKU). Kyseinen funktio arpoo satunnaisluvun, joka on vähintään 0, mutta pienempi kuin 1. Funktio arpoo uuden satunnaisluvun joka kerta kun taulukkoa päivitetään tai painetaan F9-näppäintä.

Rahanheitto (http://myy.haaga-helia.fi/~taaak/m/simu1.xlsx)

Jos haluat valjastaa RAND-funktion heittämään kolikkoa, niin voit käyttää satunnaisluku-funktiota yhdessä IF (JOS) -funktion kanssa:

=IF(RAND()<0,5;"kruunu";"klaava")

IF-funktio siis tutkii onko arvottu satunnaisluku alle 0,5 (kruunun todennäköisyys 0,5). Jos on, niin soluun asetetaan arvo "kruunu". Muutoin soluun asetetaan arvo "klaava".

Vastaavalla tavalla voit simuloida mitä tahansa binomijakautunutta satunnaisilmiötä.

Nopanheitto (http://myy.haaga-helia.fi/~taaak/m/simu2.xlsx)

Jos haluat valjastaa satunnaisluku-funktion heittämään noppaa, niin IF-funktion käyttö vaatisi useiden sisäkkäisten IF-funktioiden käyttöä. Parempi ratkaisu on käyttää aputaulukkoa

	A	B	C
1	Todennäköisyys	Kumulatiivinen	Silmäluku
2	16,7 %	0,0 %	1
3	16,7 %	16,7 %	2
4	16,7 %	33,3 %	3
5	16,7 %	50,0 %	4
6	16,7 %	66,7 %	5
7	16,7 %	83,3 %	6

Aputaulukon todennäköisyydet on laskettu kaavalla =1/6. Kumulatiiviset todennäköisyydet on saatu laskemalla A-sarakkeen todennäköisyyksiä yhteen. Huomaa, että kumulatiiviset todennäköisyydet on sijoitettu taulukkoon totutun tavan vastaisesti: esimerkiksi silmäluvun 3 kohdalla oleva kumulatiivinen todennäköisyys ilmoittaa todennäköisyyden saada alle 3.

Noppaa voidaan heittää funktion VLOOKUP (PHAKU) avulla

=VLOOKUP(RAND();B2:C7;2)

Funktio hakee arvottua satunnaislukua hakutaulukon B2:C7 ensimmäisestä sarakkeesta (VLOOKUP hakee aina hakuarvoa hakutaulukon ensimmäisestä sarakkeesta). Kun satunnaisluku tai sitä lähinnä pienempi arvo löytyy B-sarakkeesta, niin funktio hakee vastaavalta riviltä silmäluvun (argumentti 2 tarkoittaa, että funktio hakee tiedon hakutaulukon B2:C7 toisesta sarakkeesta).

Jos funktio kirjoitetaan muodossa

=VLOOKUP(RAND();$B$2:$C$7;2)

niin sitä voidaan kopioida muihinkin soluihin ja näin simuloida usean heiton sarja.

Vastaavalla tavalla voit simuloida mitä tahansa diskreettiä satunnaisilmiötä, jonka todennäköisyysjakauman voit kirjoittaa Excel-taulukkoon.

Lehtikauppiaan ongelma (http://myy.haaga-helia.fi/~taaak/m/simu3.xlsx)

Lehtikauppias on tilastoinut iltapäivälehden kysyntää. Kysyntä vaihtelee jonkin verran viikonpäivän mukaan. Lehtikauppias on laskenut tilaston perusteella kunkin viikonpäivän kysynnälle keskiarvon ja keskihajonnan. Tietyn viikonpäivän kysynnän voidaan olettaa noudattavan normaalijakaumaa.

Lehtikauppiaan täytyy ilmoittaa etukäteen tilattavien lehtien lukumäärä. Lehtikauppias pyrkii tietenkin tyydyttämään kysynnän mahdollisimman hyvin ja toisaalta hankkimaan itselleen mahdollisimman korkean voiton. Simulointimallissa simuloidaan kysynnän määrä normaalijakaumasta ja lasketaan simulointikierrosten voiton keskiarvo, keskihajonta, mediaani, minimi ja maksimi. Syöttämällä taulukkoon erilaisia tilausmääriä, voidaan seurata voiton käyttäytymistä ja löytää optimaalinen tilausmäärä.

Lehtikauppias voisi solmia nk. palautussopimuksen, jolloin myymättömistä lehdistä saa täyden hyvityksen lehtitalolta. Tässä tapauksessa lehden hinta on kauppiaalle korkeampi. Taulukossa on laskettu voiton keskiarvo, keskihajonta, mediaani, minimi ja maksimi myös lehtipalautussopimusta käytettäessä. Näin lehtikauppias voi arvioida kannattaako lehtipalautussopimusta solmia.

Lehtikauppiaan ongelmassa lehden kysyntä arvotaan normaalijakaumasta. Tämä voidaan tehdä funktiolla

=NORMINV(RAND();keskiarvo;keskihajonta)
(=NORM.JAKAUMA.KÄÄNT(SATUNNAISLUKU();keskiarvo;keskihajonta**)**.

Tuotantolinja (http://myy.haaga-helia.fi/~taaak/m/simu4.xlsx)

Tuotteet saapuvat käsiteltäviksi kahden minuutin välein. Ennen käsittelyä tuotteet tarkistetaan ja virheelliset hylätään ennen käsittelyn alkamista. Jos käsittelypiste on varattu, niin tuotteet jäävät jonoon ja otetaan käsittelyyn kukin vuorollaan. Itse käsittelyn kesto on 3 minuuttia.

Simulointimallin avulla voidaan tutkia muodostuuko käsittelypisteestä pullonkaula. Koska tuotteiden saapumisväli (2 minuuttia) ja käsittelyaika (3 minuuttia) ovat vakioita, niin ainoa kiinnostuksen kohde on hylättävien osuus käsittelypisteeseen saapuvista tuotteista. Simuloinnin lähtötietona on hylkäysprosentti, jonka arvoa voidaan vaihdella.

Tieto siitä hylätäänkö vai hyväksytäänkö tuote, saadaan arvottua funktiolla

=IF(RAND()<hylkäysprosentti;"hylkää";"hyväksy")
(=JOS(SATUNNAISLUKU()<hylkäysprosentti;"hylkää";"hyväksy"))

Asiakaspalvelupiste (http://myy.haaga-helia.fi/~taaak/m/simu5.xlsx)

Simulointimallissa uusi asiakas saapuu satunnaisesti eksponentiaalista jakaumaa noudattaen. Asiakkaan palveluaika vaihtelee normaalijakaumaa noudattaen. Simuloimalla asiakkaiden saapumista ja palvelua voidaan seurata asiakkaiden odotusaikoja ja mahdollista jononmuodostusta. Esimerkissä on simulointimalli sekä yhden palvelupisteen että kahden palvelupisteen simulointiin.

Seuraavan asiakkaan saapumisaika voidaan arpoa eksponentiaalisesta jakaumasta kaavalla (keskiarvo on asiakkaiden saapumisaikojen välin keskiarvo)

= -keskiarvo*LN(RAND()) (= -keskiarvo*LUONNLOG(SATUNNAISLUKU()))