Distribusjon av usikkerheter rundt en middelverdi – en historisk reise

v/Rune Øverland, Trainor Elsikkerhet AS

Innledning

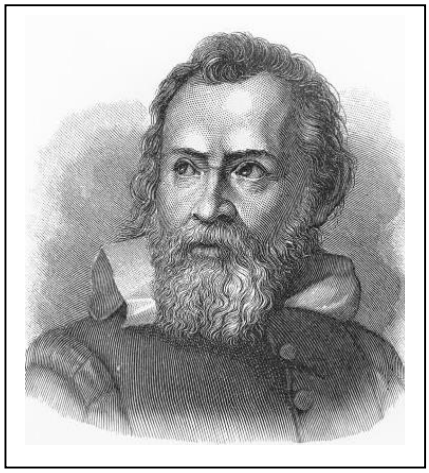
Dette er en reise som spenner over 300 år, og handler om distribusjon av usikkerheter (målevariasjoner) rundt en middelverdi. Distribusjoner er et matematisk verktøy med flere varianter, og derav flere bruksområder.

Arisoteles – Den gylne middelvei

Arisoteles (384 – 322 BC) hadde filosofiske dialoger om at det ønskede gjennomsnittet fantes mellom to ytterpunkter. Den gylne middelvei var det greske idealet for moderasjon. Den var en dyd, som stod i opposisjon til grådighet og fråtseri. Den gylne middelvei ga balanse mellom ytterlighetene.

Andre storheter som var inne på den gylne middelvei var Pytagoras, Sokrates og Platon.

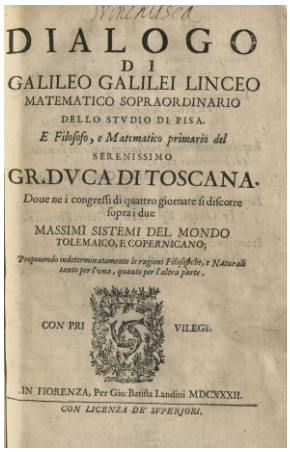
Galileo Galilei – 1632 Forløper til «Law of Error»



Galilei (1564 – 1642) var en fremragende matematiker og astronom. Galilei hørte 1609 om de første kikkertene. Rett etterpå lyktes det for ham selv å sette sammen betydelig bedre kikkerter. Han studerte nattehimmelen.

Snart oppsto det to motstridende læreteorier.

I det ptolemeiske system var jorden sentrum for universet. I det kopernikanske systemet bevegde jorden og de andre planetene seg i baner med solen som sentrum.



DIALOGO
DI
GALILEO GALILEI LINCEO
MATEMATICO SOPRAORDINARIO
DELLO STVDIO DI PISA.
E Filosofo, e Matematico primario del
SERENISSIMO
GR. DVCA DI TOSCANA.
Doue ne i congressi di quattro giornate si discorre sopra i due
MASSIMI SISTEMI DEL MONDO
TOLEMAICO, E COPERNICANO;
Proponendo indeterminatamente le ragioni Filosofiche, e Naturali tanto per l'una, quanto per l'altra parte.
CON PRI VILEGI.
IN FIORENZA, Per Gio: Batista Landini MDCXXXII.
CON LICENZA DE' SVPERIORI.

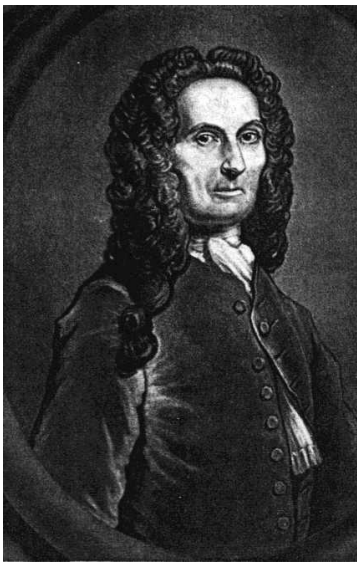
1632 (MDCXXXII): Galileo Galilei publiserte bok om tilfeldige avvik

1624 drog Galilei til Roma og forfektet med full åpenhet Copernicus' lære for den nye paven for å få opphevet inkvisisjonsdommen. Dette oppnådde han ikke.

Han utarbeidet likevel i de følgende år et verk han lenge hadde forberedt, *Dialogo sopra i Due Massimi Sistemi del Mondo* (Dialog over de to største verdenssystemer). Galilei forela dette for paven 1630, og som med visse uvesentlige tilføyelser og endringer ble tillatt trykt 1632.

Galileo Galilei hadde merket seg oppførselen til variasjoner (distribusjon av måleusikkerheter) i sine observasjoner av universet. I boken sier han om distribusjon av måleusikkerheter:

- variasjonene er uunngåelige,
- små variasjoner forekom oftere enn store,
- variasjoner kan ha samme amplitude og motsatt fortegn,
- sann verdi er der hvor vi har den høyeste konsentrasjonen av måleresultater.



Abraham deMoivre – 1733 Normalfordeling

Abraham deMoivre (1667 - 1754) var en fremragende matematiker. Tidlig ble han interessert i sannsynlighetsregning, som han fant i en bok av Huygens.

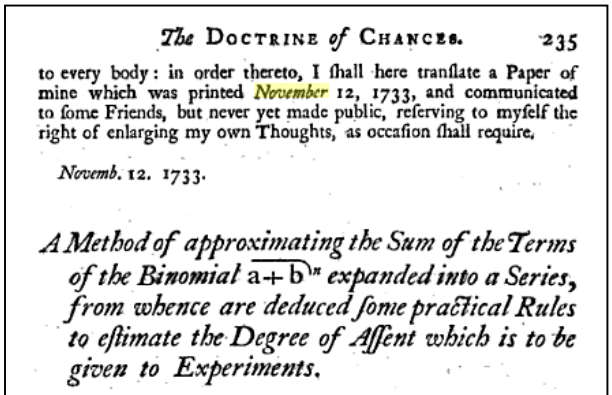
Han måtte emigrere (1692) til England, og ble raskt venn med Isaac Newton. Mange år før de traff hverandre, rundt 1665, jobbet Newton med det binomiske teorem. Et teorem som også beskrives som (Blaise) Pascals trekant (publisert 1665). Både kinesere og arabere kjente til denne tallrekken lenge før Pascal.

deMoivre's *Doctrine of Chances* (1718) er et pionérverk i sannsynlighets-regning.

Også deMoivre var fasinert i det binomiske teorem, og presenterte for noen private venner i 1733 et 7-siders dokument: «Approximatio ad summam terminorum biomii (a + b)n in seriem expansi». Dokumentet presenterte for første gang en normalfordelingskurve.

Hans egen oversettelse til engelsk, samt noen tillegg, ble inkludert i neste revisjon av boken «Doctrine of changes» som kom i 1738 (sidene 235 – 243).

Metoden bestemte sannsynligheten for at et avvik (feil) skulle forekomme for en gitt størrelse.



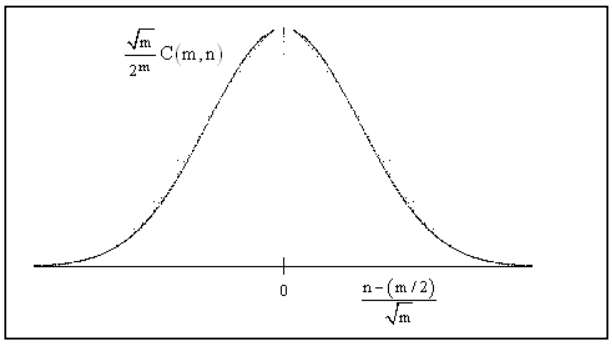
The DOCTRINE *of* CHANCES. 235

to every body: in order thereto, I ſhall here tranſlate a Paper of mine which was printed *November* 12, 1733, and communicated to ſome Friends, but never yet made public, reſerving to myſelf the right of enlarging my own Thoughts, as occaſion ſhall require.

Novemb. 12. 1733.

A Method of approximating the Sum of the Terms of the Binomial $\overline{a+b}^{n}$ *expanded into a Series, from whence are deduced ſome practical Rules to eſtimate the Degree of Aſſent which is to be given to Experiments.*

1738: Side 235 Doctrine of Changes

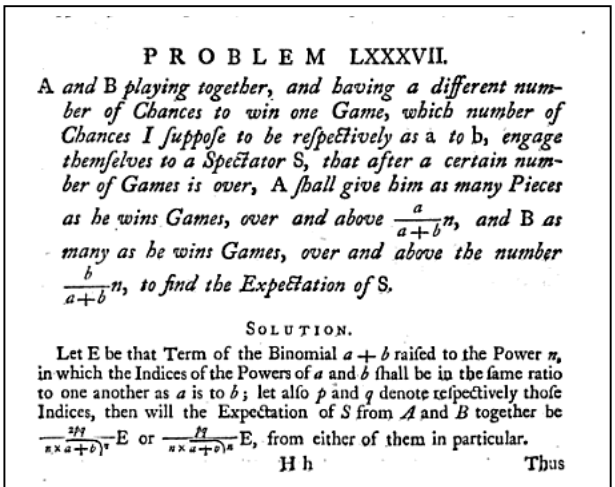


Han beskriver i «Corollary» Konsekvens 3 at ett standardavvik, som omfatter 68,2688 % av tilfellene, er bestemt av utrykket $(\frac{1}{2}n \pm \frac{1}{2}\sqrt{n})$.

Han beskriver videre i «Corollary» 4 et eksempel. Hvis man kaster kron/mynt 3600 ganger (n = 3600), er det:

$$\frac{1}{2}3600 \pm \frac{1}{2}\sqrt{3600}$$

68,2688 % sjanse for at en hendelse skulle ligge innenfor 1800 ± 30.



PROBLEM LXXXVII.

A *and* B *playing together, and having a different number of Chances to win one Game, which number of Chances I ſuppoſe to be reſpectively as* a *to* b, *engage themſelves to a Spectator* S, *that after a certain number of Games is over,* A *ſhall give him as many Pieces as he wins Games, over and above* $\frac{a}{a+b}n$, *and* B *as many as he wins Games, over and above the number* $\frac{b}{a+b}n$, *to find the Expectation of* S.

SOLUTION.

Let E be that Term of the Binomial $a+b$ raiſed to the Power n, in which the Indices of the Powers of a and b ſhall be in the ſame ratio to one another as a is to b; let alſo p and q denote reſpectively thoſe Indices, then will the Expectation of S from A and B together be $\frac{2pq}{n\times\overline{a+b}^{n}}$E or $\frac{pq}{n\times\overline{a+b}^{n}}$E, from either of them in particular.

H h Thus

1738: Oppgave 87 – tar i bruk sannsynlighetsberegninger (teori om normal distribusjon av hendelser)

COROLLARY 4.

But altho' the taking an infinite number of Experiments be not practicable, yet the preceding Conclusions may very well be applied to finite numbers, provided they be great, for Instance, if 3600 Experiments be taken, make $n = 3600$, hence $\frac{1}{2}n$ will be $= 1800$, and $\frac{1}{2}\sqrt{n} = 30$, then the Probability of the Event's neither appearing oftner than 1830 times, nor more rarely than 1770, will be 0.682688.

1733: «Corollary» 4 – eksempel på bruk av $(\frac{1}{2}n \pm \frac{1}{2}\sqrt{n})$ når n = 3600

Han hadde beskrevet (oppdaget) den *normale* sannsynlighetskurven, en kurve som ofte feilaktig tilskrives Carl Fredrich Gauss.

En annen viktig publisering deMoivre gjorde i 1733, var den første versjonen av sentralgrense-teoremet. Det tok over 200 år (1935) før teoremet ble «gjenoppdaget» for praktisk bruk.

Diskusjoner om tilfeldigheter startet sannsynligvis allerede når mennesket startet å måle. deMoivre har den første som analyserte og tolket distribusjon av usikkerheter.

Normalfordeling - et verktøy til ingen nytte?

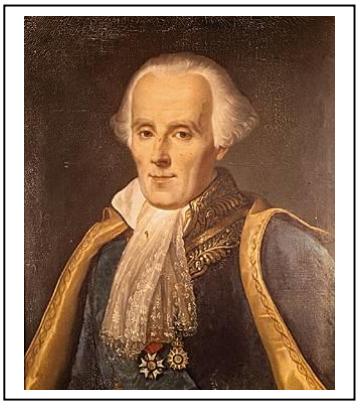
Etter at deMoivre hadde beskrevet sin teori om normalfordeling i 1733 for sine nærmeste venner, og deretter publiserte dette i 1738, var det få eller ingen som fatter interesse i dette verktøyet.

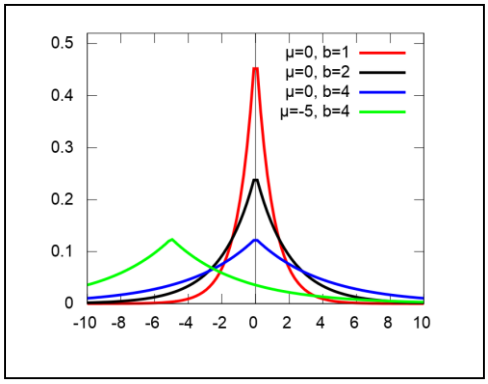
17-hundretallet var først og fremst i planeters og kometers navn.

deMoivre's normaldistribusjon var «bare» et spesialtilfelle av binomisk distribusjon (spill-teori) som et resultat at tilfeldige operasjoner som å kaste kron/mynt, kaste terninger eller trekke kort fra en kortstokk. Hans oppdagelse av normaldistribusjon fattet ikke mye interesse blant astronomer.

Pierre-Simon LaPlace – 1774 «Laws of Errors»

LaPlace (1749 – 1827) publiserte 1774 første lov om distribusjon av usikkerheter. Han oppdaget at frekvensen av usikkerheter kunne uttrykkes ved hjelp av eksponentiell-funksjoner.





LaPlace satte to speilvendte eksponentialfunksjoner inn mot hverandre for å lage én ny kurve (fordeling av usikkerheter).

Kurvene ble et alternativ til normalfordelingskurven til deMoivre fra 1733.

LaPlace publiserte 1778 andre lov om distribusjon av usikkerheter.

Carl Friedrich Gauss – 1809 Standardisert normalfordelingskurve

Carl Friedrich Gauss

Carl Friedrich Gauss (1777 – 1855) publiserte 1809 Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium.

Porro facile perspicitur, k necessario negatiuam esse debere, quo Ω reuera fieri possit maximum, quamobrem statuemus $\frac{1}{2}k = -hh$; et quum per theorema elegans primo ab ill. Laplace inuentum, integrale $\int e^{-hh\Delta\Delta} d\Delta$, a $\Delta = -\infty$ vsque ad $\Delta = +\infty$, fiat $= \frac{\sqrt{\pi}}{h}$, (denotando per π semicircumferentiam circuli cuius radius 1), functio nostra fiet

$$\varphi\Delta = \frac{h}{\sqrt{\pi}} e^{-hh\Delta\Delta}$$

1809 side 212: Gauss publiserer likning for standardisert normalfordeling

Her publiserer han blant annet teorien om (standardisert) normalfordelingskurve.

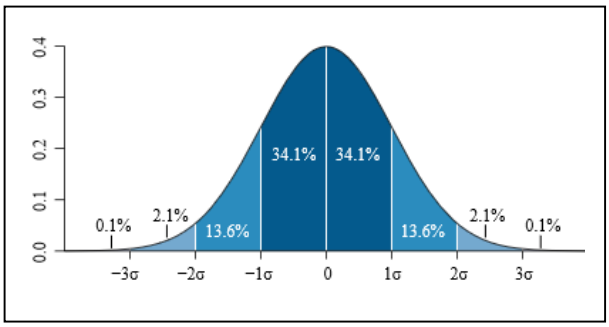
Følgende likning fins på side 212:

$\varphi\Delta = \frac{h}{\sqrt{\pi}} e^{-hh\Delta\Delta}$, hvor Δ er «error» sentrert om 0, og *h* beskrives som «modulus».

Funksjon som bestod av tre parametere, ble senere omskrevet til;

$$f(x,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left[\frac{(x-\mu)^2}{2\sigma^2}\right]}$$

hvor μ er sentralverdien, σ er standardavviket. (x- μ) uttrykker avviket fra den observerte verdien og sentralverdien. Gauss viser at observerte variasjoner distribueres etter ovennevnte kurve.



For Gausskurven gjelder at ett standardavvik omfatter 68,2689492 % av tilfellene.

Som tidligere nevnt; allerede i 1733 jobbet Abraham deMoivre (med litt assistanse fra Stirling) med normaldistribusjon basert på binomial distribusjon, som kunne skrives:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\left[\frac{(x)^2}{2}\right]}$$

For Moivre's likning var ett standardavvik 68,2688 %. Det skilte altså bare noen milliontedeler mellom de to kurvene! (deMoivre 0,68268**8** og Gauss 0,68268**9**)

For å skille mellom de to normaldistribusjonene, omtales ofte normalfordelingskurven til Gauss for Gauss-kurven.

Pierre-Simon LaPlace – 1810/1812 en gang til . . .

LaPlace avdekker noen svakheter i Gauss sitt arbeid med normalfordelingskurven (fra 1809), og publiserer 1810 en forbedret versjon i sitt supplement *Mémoire sur les approximations des formules qui sont fonctions de très grands nombres et sur leur application aux probabilités.* Han kommer også tilbake til sine arbeider fra 1774 og 1778.

Tidligere var LaPlace mest interessert i fordelinger for Games»/spill. Nå han konsentrerte han seg om fordelinger for sosiale distribusjoner.

1812 utgir LaPlace *Théorie analytique des probabilités.* Her forbedrer han deMoivre sin likning fra 1733 om normalfordelingskurven. Den kalles heretter for deMoivre-LaPlace sitt teorem.

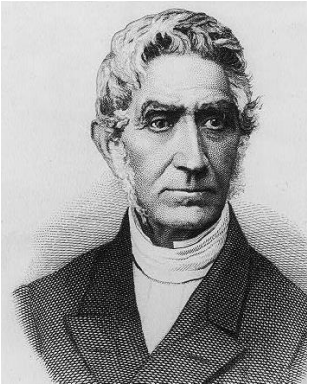
Friedrich Wilhelm Bessel – 1815 'Far' til usikkerhetsintervallet

Bessel (1784 - 1846) definerer i 1815 intervallet for sannsynlig feil (wahrscheinlichkeit Fehler) som avstanden mellom median (gjennomsnittsverdi) og en fjerdedel i en normalfordeling. Dette tilsvarer 0,6745 standardavvik. Dette var uttrykket for variasjon inntil det ble avløst av uttrykket standardavvik (Pearson 1893).

Bessel definerte i 1815 et intervall der man med 50 % sannsynlighet at en hendelse skulle skje (ved ± 0,6745 standardavvik) basert på Gausskurven.

1815: Friedrich Bessel definerer variasjon

Adolphe Quételet – 1846 Fra teori til praksis?



Adolphe Quételet

Lambert Adolphe Jacques Quételet (1796 – 1874) er mest kjent for sine studier innen statistikk. På en systematisk måte innhentet han statistikk om befolkninger.

På starten av 1800-tallet hadde det vært store omveltninger i befolkningen. Man hadde behov for å forstå dynamikken i samfunnet. Quételet var opptatt av gjennomsnitt-mennesket, og han søkte bevis på ryddighet av samfunnet.

Quételet tok normalfordelingskurven som bevis for eksistensen av orden i den tilsynelatende kaotiske, sosiale verden.

Quetelet, far til kvantitative samfunnsvitenskap, var den første til å hevde at normalfordelingskurven kunne brukes på tilfeldige variasjoner, også til fordelinger av sosiale fenomener.

De tidlige samfunnsvitere (også Quételet) hadde dårlig forståelse av de matematiske forutsetningene som lå i normalfordelingskurven. Noen hadde den tiltro og tillit til normalfordelingskurven at all mulig samfunnsvitenskap skulle passe inn i den!

På den annen side, matematikerne som deMoivre og Gauss følte ikke at det var deres domene å sjekke hvorvidt den empiriske verden passet inn i deres normalfordelingskurve!

I 1832 utarbeidet Quételet hva som omtales som «Quételet -indeksen», mer kjent i dag som «Body Mass Index» - BMI – kroppsmasseindeksen. Han studerte også vekt og lengde på nyfødte barn.

QUETELET'S TWO ATTEMPTS 207

MESURES de la POITRINE.	NOMBRE d'hommes.	NOMBRE PROPORTIONNEL.	PROBABILITÉ d'après L'OBSERVATION.	RANG dans LA TABLE.	RANG d'après le CALCUL.	PROBABILITÉ d'après LA TABLE.	NOMBRE D'OBSERVATIONS calculé.
Pouces.							
33	3	5	0,5000			0,5000	7
34	18	31	0,4993	52	50	0,4993	29
35	81	141	0,4964	42,5	42,5	0,4964	110
36	185	322	0,4823	33,5	34,5	0,4854	323
37	420	732	0,4501	26,0	26,5	0,4531	732
38	749	1305	0,3769	18,0	18,5	0,3799	1333
39	1073	1867	0,2464	10,5	10,5	0,2466	1838
			0,0597	2,5	2,5	0,0628	
40	1079	1882	0,1285	5,5	5,5	0,1359	1987
41	934	1628	0,2913	13	13,5	0,3034	1675
42	658	1148	0,4061	21	21,5	0,4130	1096
43	370	645	0,4706	30	29,5	0,4690	560
44	92	160	0,4866	35	37,5	0,4911	221
45	50	87	0,4953	41	45,5	0,4980	69
46	21	38	0,4991	49,5	53,5	0,4996	16
47	4	7	0,4998	56	61,8	0,4999	3
48	1	2	0,5000			0,5000	1
	5738	1,0000					1,0000

1846: Tabulerte verdier fra 1817

I 1846 publiserte han brystmål på 5738 skotske soldater. (Det var egentlig 5732 soldater, men Quételet hadde summert litt feil fra de skotske journalene av 1817).

Menneskelige karakteristikker ble for første gang presentert. Kolonne 1 har brystmålene. Kolonne 2 gir informasjon om hvor mange av soldatene som hadde dette brystmålet. Kolonne 3 var normalisert (skalert) pr 10.000 soldater.

Quételet organiserte den første internasjonale statistikk-konferanse i 1853.

Auguste Bravais – 1846 Fra teori til praksis (= korrelasjon)!?

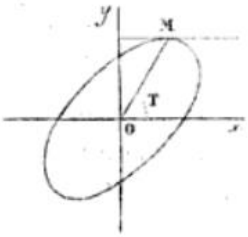
Auguste Bravais (1811 – 1863) ga ut i 1846 «Analyse mathématique sur les probabilités des erreurs de situation d'un point». Dette er vurdert av historikere til å være den første boken om korrelasjonsteori.

1846: Auguste Bravais

Her diskuterer Bravais grunnleggende teoremer om korrelasjon. Det var behov for å sjekke i hvilken grad empiriske data/observasjoner stemte overens med normalfordelingskurven.

DES ERREURS DE SITUATION D'UN POINT. 19

Le rayon vecteur, correspondant à ces valeurs de x et de y, sera le demi-diamètre conjugué de l'axe des x, et si T représente l'angle de ces deux diamètres, $\frac{y}{\sin.T} = \frac{1}{\sin.T}\sqrt{\frac{ab}{ab-r^2}}$ sera la grandeur OM de ce rayon vecteur (voyez figure ci-dessous).



På bokens side 19 tegner Bravias sin berømte illustrasjon om korrelasjon

Bravais var først og fremst opptatt av arealet av ellipser, og ikke av den rette linjen OM (se figur).

Bravais finner vi hva senere omtaler som regresjonslinjen (linjen OM).

Francis Galton – 1869 Fra teori til praksis («Law of Errors»)!

Francis Galton (1822 – 1911), fetter av Charles Darwin, var interessert i alt og alle. Det sies at hans IQ var over 200. Han studerte blant annet matematikk og medisin. Hans slektskap til Darwin gjorde at han ble sterkt influert av utviklingsteorien. Charles Darwin utga i 1859 «The Origin of Species». Galton lurte på om den teoretiske normalfordelingskurven gjenspeilte naturen. Inspirert av sin fetter Darwin publiserer Galton 1869 sitt «Hereditary Genius».

Francis Galton

Galton tar blant annet utgangspunkt i tallmaterialet fra Quételet.

TABLE BY QUETELET.

Grade or Rank of the Group.	M Probability of Drawing each Group.	N Sum of the Probabilities, commencing; at the most probable Group.	Number of the Grade.	M Probability of Drawing each Group.	N Sum of the Probabilitie commencing at the mo: probable Group.
1	.025225	.025225	41	.0009458	.495278
2	.025124	.050349	42	.0008024	.496081
3	.024924	.075273	43	.0006781	.496759
4	.024627	.099900	44	.0005707	.497329
5	.024236	.124136	45	.0004784	.497808
6	.023756	.147892	46	.0003994	.498207
7	.023193	.171085	47	.0003321	.498539
8	.022552	.193637	48	.0002750	.498814
9	.021842	.215479	49	.0002268	.499041
10	.021069	.236548	50	.0001863	.499227
11	.020243	.256791	51	.0001525	.499380
12	.019372	.276163	52	.0001243	.499504
13	.018464	.294627	53	.0001008	.499605
14	.017528	.312155	54	.0000815	.499686
15	.016573	.338728	55	.0000656	.499752
16	.015608	.344335	56	.0000526	.499804
17	.014640	.358975	57	.0000421	.499847
18	.013677	.372652	58	.0000334	.499880
19	.012726	.385378	59	.0000265	.499906
20	.011794	.397172	60	.0000209	.499927
21	.010887	.408060	61	.0000164	.499944
22	.019998	.418070	62	.0000128	.499957
23	.009106	.427236	63	.0000100	.499967
24	.008360	.435595	64	.0000077	.499974
25	.007594	.443189	65	.0000000	.499980
26	.006871	.450060	66	.0000046	.499985
27	.006191	.456251	67	.0000035	.499988
28	.005557	.461809	68	.0000027	.4999912
29	.004968	.466776	69	.0000021	.4999933
30	.004423	.471199	70	.0000016	.4993948
31	.003922	.475122	71	.0000012	.4999960
32	.003464	.478586	72	.0000009	.4999969
33	.003047	.481633	73	.0000007	.4999976
34	.002670	.484304	74	.0000005	.4999981
35	.002330	.486634	75	.0000004	.4999984
36	.002025	.488659	76	.0000003	.4999987
37	.001753	.490412	77	.0000002	.4999989
38	.001512	.491924	78	.00000014	.4999990
39	.001208	.493222	79	.00000011	.4999991
40	.001110	.494332	80	.00000004	.4999991

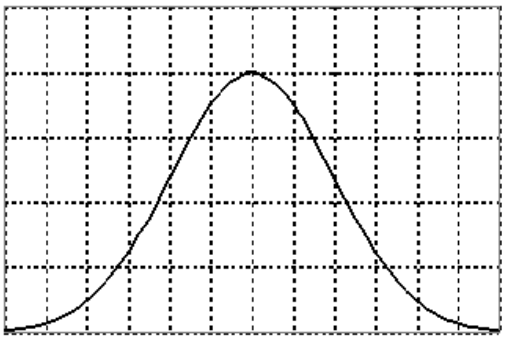
1869: Hereditary Genius side 378

Galton kommenterer følgende:
"I know of scarcely anything so apt to impress the imagination as the wonderful form of cosmic order expressed by the "law of error".

A savage, if he could understand it, would worship it as a god. It reigns with severity in complete self-effacement amidst the wildest confusion. The huger the mob and the greater the anarchy, the more perfect is its sway.

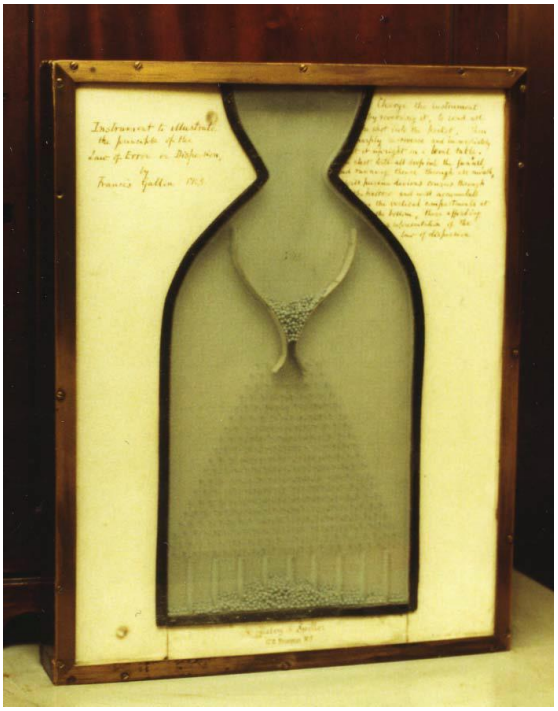
Let a large sample of chaotic elements be taken and marshalled in order of their magnitudes, and then, however wildly irregular they appeared, an unexpected and most beautiful form of regularity proves to have been present all along.

Arrange statures side by side in order of their magnitudes, and the tops of the marchalled row will form a beautifully flowing curve of invariable proportions; each man will find, as it were, a pre-ordained niche, just of the right height to fit him, and if the class-places and statures of any two men in the row are known, the stature that will be found at every other place, except towards the extreme ends, can be predicted with precision."



Her presenterer Galton effekten av normal-distribusjon av «kosmisk kaos». Begrepet «normal distribution» bruker Galton i 1875.

Hans arbeider på 1870-tallet leder oss inn i regresjonsteorier og bruk av minste kvadraters metode.



1873: Galtons Quincunx-maskin

I hans Quincunx-maskin slapp han kuler i senter. Kulene fikk en distribusjon (normalfordeling) slik som den matematiske modellen til Gauss. Maskinen fungerte perfekt!

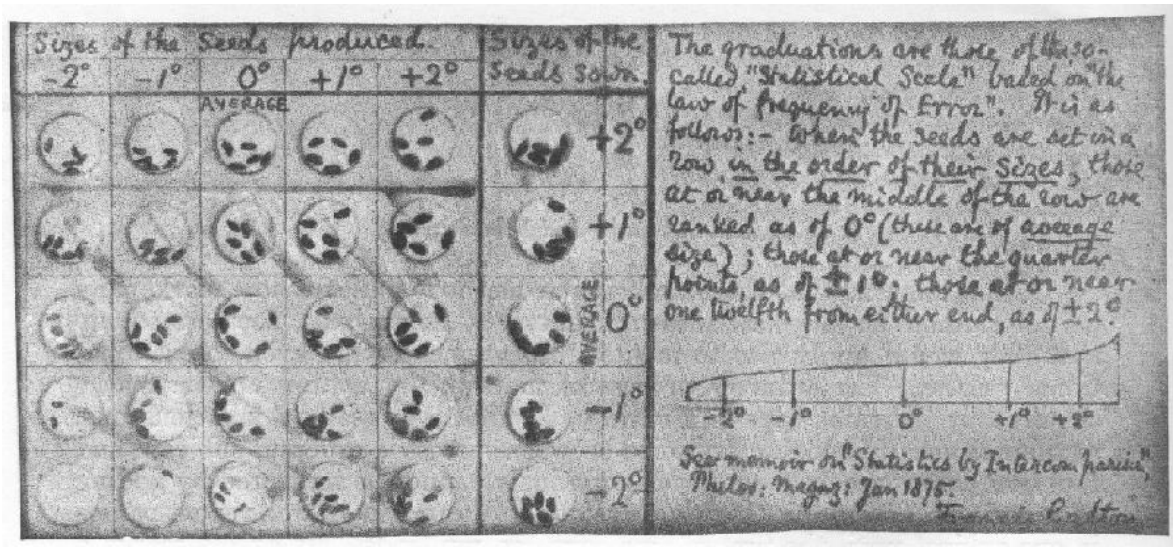
Han brukte også denne Quincrux-maskien til å demonstrere sentralgrenseteoremet (deMoivre 1733/LaPlace 1778) for å vise at normaldistribusjon er tilnærmet lik binominal distribusjon.

Regresjon - Sukkererter - 1877

Galton gjorde mange statistiske studier. Han søkte bevis for sine teoremer.

Dette leder oss til sukkererter. I 1875 distribuerte han ulike pakker med sukkererter til 7 venner. Hver pakke hadde sukkererter av samme vekt, mens pakkene var ulike.

Senere plottet han resultatet, og så på sammenhengen mellom «mor» og «barn». I 1877 plottet Galton vektene av datterfrø mot vektene av morfrø. Galton fant at gjennomsnittsvekt av datterfrø fra en bestemt størrelse av morfrø omtrent beskrev en rett linje med positiv helning mindre enn 1,0.



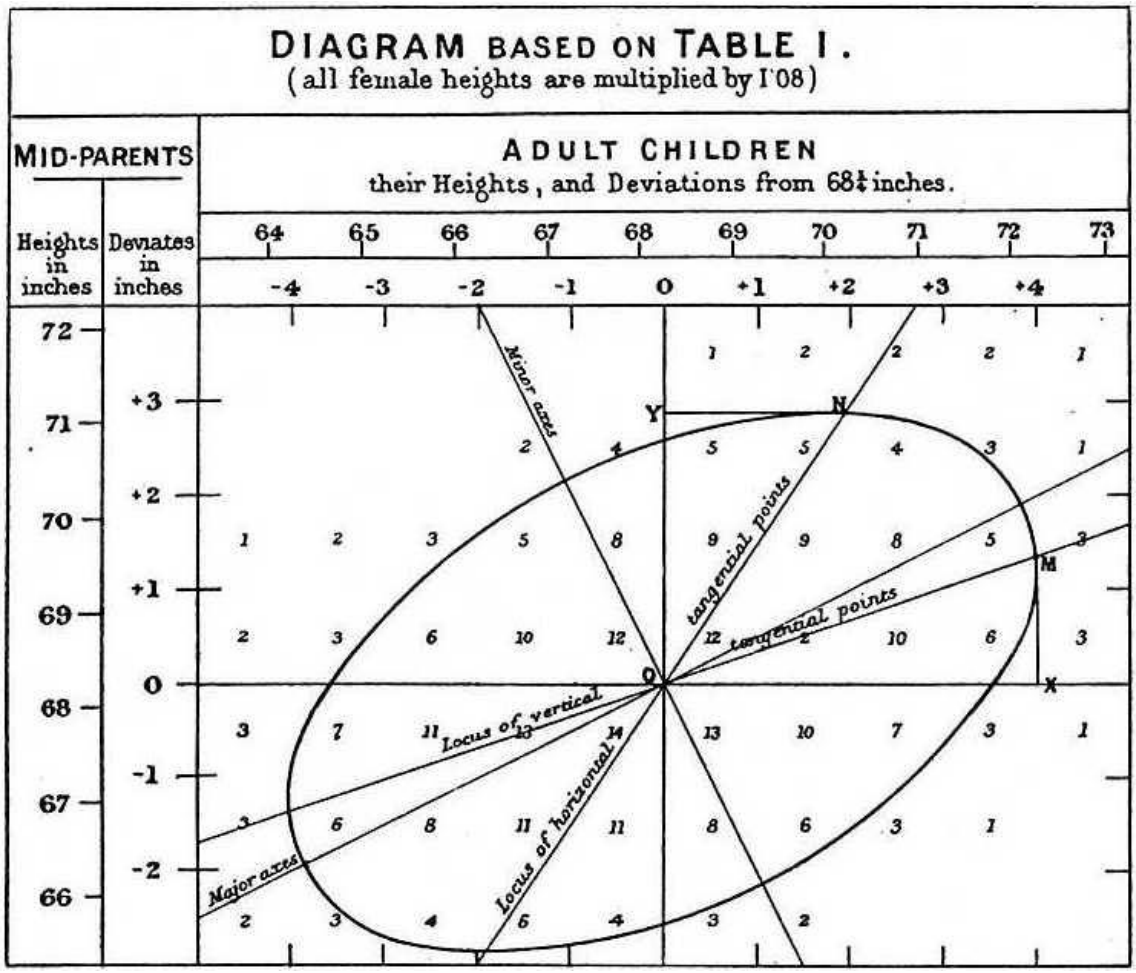
1877: Galtons statistiske analyse av sukkerertene

Galton jobbet nå mer aktivt med regresjonsteori enn tidligere; uten at han kunne sette ord på arbeidet.

Regresjonsmodellen; 1886 - gjennombruddet

I 1886 publiserte Galton sine observasjoner fra en eksperimentell test av høyden på familier. 205 foreldre og 928 barn inngikk i statistikken. Han ville finne den statistiske sammenhengen mellom høyden til barn og deres foreldre.

Kvinners høyde ble multiplisert med en faktor på 1,08 for å samsvare med høyden på menn.



1886: Galtons regresjonslinjer (2 stykk) inntegnet i ellipse

Senter av ellipsen (O) samsvarer med gjennomsnittshøyden av foreldrene (horisontal akse), og barna (vertikal akse).

Galton tegnet inn den vertikale linjen XM, hvor M er punktet lengst til høyre på ellipsen. Dette dannet linjen OM – «Locus of vertical». Denne linjen beskriver den mest sannsynlige fedrehøyden for en gitt middels barnehøyde. Galton tegnet også inn en linje YN, hvor N er det øverste punktet på ellipsen. Han tegnet inn tangenten ON – «Locus of horizontal». Denne linjen beskriver den mest sannsynlige barnehøyden for en gitt middels fedrehøyde.

Linjen «Major axes» tegnes inn på ellipsens maksimale avstand, mens «Minor axes» er tegnet inn vinkelrett på denne (på ellipsens smaleste punkt).

Ellipsen hadde mange fellestrekk med ellipsen til Bravais fra 1846.

I 1888 definerte Galton begrepet «Co-relation», et uttrykk han senere kalte «Correlation».

Kjært barn har mange navn - «Bell»-kurven -1872

Normalfordelingskurven har senere gått under ulike navn. Ett av de er «Bell»-kurven. Dette fordi normalfordelingskurven kan se ut som en (kirke)klokke.

Esprit Jouffret (1837 – 1904) var den første som karakteriserte normalfordelingskurven som en klokke. Det skjedde i 1872. Han skriver «La surface S, en forme de cloche» i sitt *Etude sur l'effet utile du tir*.

Galton brukte i 1876 uttrykket «Bell shaped Curve».

Esprit Jouffret

Karl Pearson – 1893 Standardavvik

Francis Galton var en meget nær venn og kollega av Karl Pearson (1857 – 1936). Carl (etter hvert kjent som Karl), jobbet også med andre storheter som Albert Einstein.

I 1893 definerte Pearson standardavviket, slik vi kjenner det i dag:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{X})^2}$$

Vi trekker kvadratroten av summen av avvikene mellom observasjoner (x_i) og gjennomsnittet av observasjonene ($\bar{X}$), som deles på antall observasjoner (N).

1909: Karl Pearson (venstre) sammen med Francis Galton

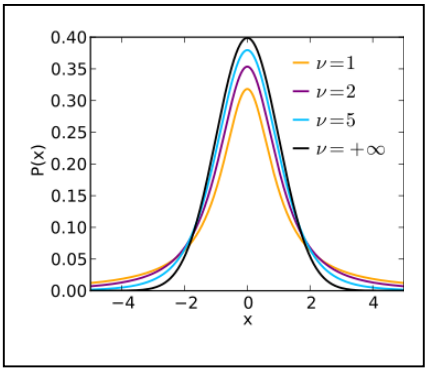
William Sealey Gosset - 1908 Student t-fordeling

William S. Gosset «Student»

Gosset (1876 – 1937) jobbet som kjemiker på et bryggeri (Guinness). Han jobbet med å forbedre kvaliteten på ølet, hvor bygg inngikk som råvare. Gosset var god venn med Karl Pearson.

Guinness forbød sine ansatte å publisere sine statistiske arbeider. Han publiserte (1908) derfor under pseudonymet Student sin Students t-fordeling i journalen Biometrika.

Han definerte t-fordeling for å håndtere små fordelinger (små stikkprøver) av bygg for kvalitetskontroll i bryggerivirksomheten. Gosset observerte at når han tok små fordelinger (små stikkprøver) av bygg, stemte fordelingen ikke helt med normalfordelingskurven. Men, desto større fordelinger (større stikkprøver) av bygg han tok, desto mer nærmet fordelingen seg normalfordelingskurven.



Han publiserte derfor en faktor (t-fordeling) som korrigerer for fordelinger av ulike størrelser slik at disse passer med normalfordelingskurven. Han fant at når antall 'individer' var 30 eller flere, stemte fordelingen godt overens med normalfordelingskurven.

Walter Andrew Shewhart – 1931: 95 % usikkerhetsgrense

Walter A Shewhart

Shewhart (1891 – 1967) er kjent som grunnlegger av produktstyring ved hjelp av statistikk. Grunnlaget var normalfordelingskurven og etablering av usikkerhetsgrenser.

Han startet å arbeide i Bell Telephone Laboratories i 1925. Han oppsummerte sine arbeider med statistikkstyring i 1931 i hans bok *Economic Control of Quality of Manufactured Product*.

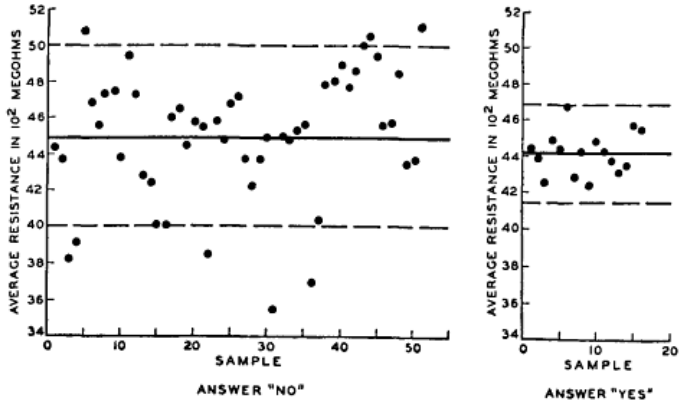


Fig. 7.—Should These Variations be Left to Chance?

Usikkerhetsgrensene; UCL – Upper Control Limit og LCL – Lower Control Limit, ble et intervall hvor produkt-kvaliteten skulle ligge.

Dagens industristandard bygger på at kontrollgrensene (usikkerhetsgrensene) skulle plasseres slik at 95 % av tilfellene (produktene) skulle ligge innenfor grenseverdiene. En 95 % mengde tilsvarer at variasjonen er beskrevet med 2 ganger standardavviket. I enkelte tilfeller er en akseptgrense på 95 % uforenlig. Det betyr at 5 % av produktene kastes. Alternativt kan usikkerhetsgrensene settes ved 3 ganger standardavvik, hvilket vil inkludere 99,7 % av produktene.

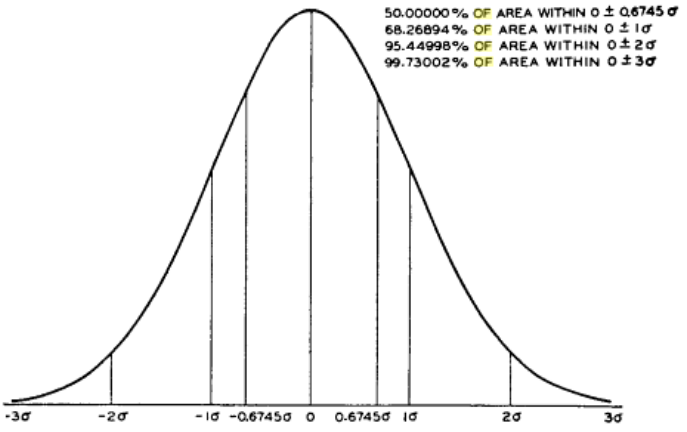


Fig. 33.—Normal Law Curve.

Du har nå, med 95 % usikkerhet sannsynlighet, fått et bedre grunnlag til å forstå måleskriftens § 8 som sier: «Det skal utarbeides usikkerhetsanalyse for målesystemet innenfor 95 % konfidensnivå."

Med vennlig hilsen

Trainor Elsikkerhet AS

Rune Øverland
Senioringeniør

Tønsberg mai 2015