Innledning til STK1100

(dekker også deler av notatet om *Sannsynlighetsbegrepet* og avsnittene 1.1 og 1.2 i læreboka)

Januar 2008

Ørnulf Borgan
(borgan@math.uio.no)
Matematisk institutt
Universitetet i Oslo

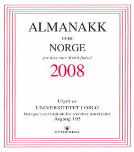
Deterministiske fenomener

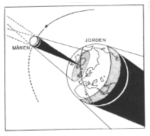
Almanakk for Norge viser:
- når det er fullmåne
- når det er soloppgang og solnedgang

Det er mulig siden astronomene kan gi en matematisk beskrivelse av himmellegemenes bevegelser

De kan også regne ut når vi vil få solformørkelse

Fullmåne, soloppgang/nedgang og solformørkelse kan forutsies. De er *determinstiske* fenomener





Fra Store Norske Leksikon

Stokastiske forsøk

Vi vet ikke hva resultatet vil bli:
- når vi kaster en terning
- i neste ukes lottotrekning

LOTTO

Terningkast og lottotrekning er eksempler på *stokastiske forsøk* (tilfeldige forsøk)

Det er også et stokastisk forsøk å se hvilket kjønn et nyfødt barn har

Kjennetegnet på et stokastisk forsøk er at vi *ikke* kan forutsi resultatet.

Sannsynlighetsregning er matematikken for stokastiske forsøk

Litt historikk

Menneskene har hatt terninger i tusenvis av år:
- Astragalus: lagd av en liten knokkel fra foten til en sau eller hund
- Sekssidet terning lagd for eksempel av bein, stein eller bronse

(www.historicgames.com)

(www.gambletribune.org)

Terningkast og loddtrekning ble brukt for å komme fram til riktig avgjørelse på et vanskelig problem. Avgjørelsen ble da overlatt til høyere makter

Apg. 1, 23-26:

To menn ble kalt fram, Josef Barsabbas med tilnavnet Justus, og Mattias. Så bad de: «Herre, du som kjenner alles hjerter, vis oss hvem av disse to du har utvalgt til å ha den tjeneste og det apostelembete som Judas forlot for å gå til sitt sted.» De kastet lodd mellom dem, og loddet falt på Mattias. Fra nå av ble han regnet som apostel sammen med de elleve.

Torstein Frode sier at det var en bygd på Hisingen som snart hadde fulgt med Norge og snart med Gøtaland. Nå avtalte kongene med hverandre at de skulle kaste lodd om hvem som skulle eie den; de skulle kaste terninger, og den som fikk størst, skulle ha den. Sveakongen kastet to seksere, og så sa han at kong Olav trengte ikke kaste. Han ristet terningene i hånden og sa: «Det er to seksere på terningen ennå, og det er ingen sak for Gud min herre å la dem komme opp.» Han kastet, og det kom opp to seksere. Så kastet Olav sveakonge, og det ble to seksere igjen. Så kastet Olav Norges konge, og da kom det opp seks på den ene, men den andre gikk i stykker, så det kom opp sju på den. Da fikk han bygda. Vi har ikke hørt noe annet å fortelle fra dette møtet. Kongene skiltes som forlikte.

Kongene kaster terninger om en bygd på Hisingen.

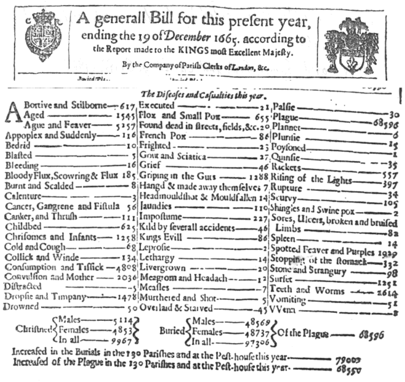
Blaise Pascal (1623-1662)

Pierre de Fermat (1601-1665)

Det historiske gjennombruddet for sannsynlighetsregningen kom først i 1654 i en brevveksling mellom Pascal og Fermat om noen problemer knyttet til spill

(Bilder fra www.york.ac.uk/depts/maths/histstat)

Etter hvert rettet interessen seg også mot andre områder enn spill. Blant annet ble livsforsikring vanligere blant de velstående på 1700-tallet. Og da måtte en kunne beregne overlevelses-sannsynligheter med utgangs-punkt i empiriske data



A generall Bill for this preſent year, ending the 19 of *December* 1665. according to the Report made to the KINGS moſt Excellent Majeſty.

By the Company of Pariſh Clerks of *London*, &c.

The Diſeaſes and Caſualties this year.

Abortive and Stilborne	617	Executed	21	Palſie	30
Aged	1545	Flox and Small Pox	655	Plague	68596
Ague and Feaver	5257	Found dead in ſtreets, fields, &c.	20	Plannet	6
Appoplex and Suddenly	116	French Pox	86	Pluriſie	15
Bedrid	10	Frighted	23	Poyſoned	1
Blaſted	5	Gout and Sciatica	27	Quinſie	35
Bleeding	16	Grief	46	Rickets	557
Bloody Flux, Scowring & Flux	185	Griping in the Guts	1288	Riſing of the Lights	397
Burnt and Scalded	8	Hangd & made away themſelves	7	Rupture	34
Calenture	3	Headmouldſhot & Mouldfallen	14	Scurvy	105
Cancer, Gangrene and Fiſtula	56	Jaundies	110	Shingles and Swine pox	2
Canker, and Thruſh	111	Impoſtume	227	Sores, Ulcers, broken and bruiſed Limbs	82
Childbed	625	Kild by ſeverall accidents	46	Spleen	14
Chriſomes and Infants	1258	Kings Evill	86	Spotted Feaver and Purples	1929
Cold and Cough	68	Leproſie	2	Stopping of the ſtomack	332
Collick and Winde	134	Lethargy	14	Stone and Strangury	98
Conſumption and Tiſſick	4808	Livergrown	20	Surfet	1251
Convulſion and Mother	2036	Meagrom and Headach	12	Teeth and Worms	2614
Diſtracted	5	Meaſles	7	Vomiting	51
Dropſie and Timpany	1478	Murthered and Shot	5	VVenn	8
Drowned	50	Overlaid & Starved	45		

Chriſtned: Males 5114, Females 4853, In all 9967

Buried: Males 48569, Females 48737, In all 97306; Of the Plague 68596

Increaſed in the Burials in the 130 Pariſhes and at the Peſt-houſe this year 79009

Increaſed of the Plague in the 130 Pariſhes and at the Peſt-houſe this year 68590

(www.york.ac.uk/depts/maths/histstat)

7

I dag brukes stokastiske modeller og sannsynlighetsregning på en rekke områder. For eksempel innen:

- Forsikring og finans
- Medisin og genetikk
- Signal- og bildebehandling
- Forvaltning av dyrepopulasjoner

Temaet for STK1100 er grunnleggende sannsynlighetsregning og stokastisk modellering

Sannsynlighetsregningen er også grunnlaget for statistiske metoder (STK1110)

8

Utfall og utfallsrom

Generelt er resultatet av et stokastisk forsøk *ikke* gitt på forhånd. Men vi kan angi de mulige resultatene, eller *utfallene* for forsøket

Mengden av alle utfall kaller vi *utfallsrommet* Ω

Eksempel 1: Når vi kaster en terning vet vi *ikke* hvor mange øyne vi vil få. Men vi vet at antall øyne vil bli 1, 2, 3, 4, 5 eller 6

Utfallsrommet er $\Omega = \{1, 2, 3, 4, 5, 6\}$

9

Eksempel 2: Kast en terning til første gang du får en sekser.

Utfallsrom: $\Omega = \{1, 2, 3, 4, \ldots\}$

Eksempel 3: Registrer vekten til en nyfødt jente.

Kan ikke her liste opp alle mulige fødslesvekter.

Utfallsrommet er er intervall på tallinja, for eksempel $\Omega = (0, 6)$

10

I eksempel 1 er utfallsrommet *endelig,* mens det er *tellbart uendelig* i eksempel 2

En fellesbetegnelse for disse to situasjonene er at utfallsrommet er *diskret*

I eksempel 3 er utfallsrommet *kontinuerlig*

11

Begivenheter

Vi vil ofte være interessert i et resultat av et forsøk som svarer til flere utfall

Et resultat av et forsøk som svarer til ett eller flere utfall kaller vi en *begivenhet* (hendelse)

Eksempel 1 (forts): Ved terningkast kan vi være interessert i om vi får *minst* fem øyne

Begivenheten A = "minst fem øyne" kan vi skrive $A = \{5, 6\}$

12

Eksempel 2 (forts): Kast terning til første sekser

A = "høyst tre kast" = {1, 2, 3}

B = "minst fem kast" = {5, 6, 7, … }

Eksempel 3 (forts): Registrer vekten til en nyfødt jente.

A = "veier mellom 3 kg og 4 kg" = (3 , 4)

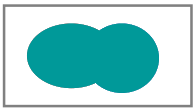
13

Union og snitt

La *A* og *B* være to begivenheter ved et forsøk
Ut fra disse kan vi lage to nye begivenheter:

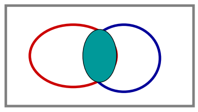
$A \cup B$

omfatter alle utfall som er med i *A* eller i *B* eller i begge ("*A* union *B*")



$A \cap B$

omfatter alle utfall som er med i både *A* og *B* ("*A* snitt *B*")



Hvis det ikke fins noen utfall som er med i både *A* og *B* er begivenhetene *disjunkte*

14

Komplementære begivenheter

La *A* være en begivenhet ved et forsøk

Det at *A* *ikke* inntreffer er en begivenhet vi kaller "ikke *A*" og skriver A^c

A og A^c er *komplementære* begivenheter

Merk at *A* og A^c er *disjunkte*

15

Det klassiske sannsynlighetsbegrepet

De første arbeidene om sannsynlighet var om problemer i spill, og sannsynlighetsbegrepet de brukte var tilpasset dette

Et enkelt eksempel illustrerer tankegangen

Vi kaster en terning

Hva er sannsynligheten for at vi får minst fem øyne?

16

Alle sidene på terningen har *samme sannsynlighet* for å vende opp når terningen kastes

Sannsynligheten er lik 1 for at en eller annen side vil vende opp

Derfor er sannsynligheten 1/6 for hver av de seks utfallene 1, 2, 3, 4, 5 og 6

Begivenheten "minst fem øyne" består av de to utfallene 5 og 6

Sannsynligheten for minst fem øyne er derfor 2/6 = 1/3

17

Generell formulering av argumentet:

Et stokastisk forsøk har *N* utfall
Det er de *mulige utfallene* for forsøket

Vi antar at de *N* utfallene er *like sannsynlige*

Da her hvert utfall sannsynlighet 1/*N*

En begivenhet *A* består av *n* utfall
Det er de *gunstige utfallene* for begivenheten *A*

Sannsynligheten for begivenheten *A* er

$$\boxed{P(A) = \frac{n}{N} = \frac{\text{antall gunstige utfall}}{\text{antall mulige utfall}}}$$

18

Sannsynlighet og relativ frekvens

Hva betyr det at sannsynligheten er 1/6 for sekser i et terningkast?

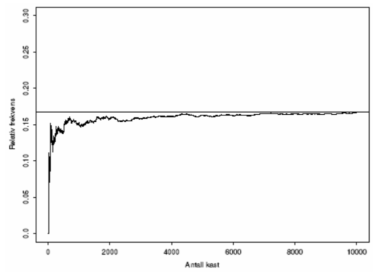
Vi kan ikke forutsi resultatet av ett kast, men vi ser et mønster når terningen kastes mange ganger

Kaster først en terning 10 ganger:

Relativ frekvens av seksere er 1/10 = 0.10

Kaster så en terning om og om igjen. Etter *k* kast er den relative frekvensen av seksere:

$$r_k(6) = \frac{\text{antall seksere i } k \text{ kast}}{k}$$



Merk at den relative frekvensen nærmer seg 1/6=0.167 når *k* blir stor

Hva er sannsynligheten for at et nyfødt barn er en jente?

År Årsgj.sn	Levendefødte I alt	Gutter	Jenter
1951-55	62478	32182	30296
1956-60	63021	32374	30647
1961-65	63989	32992	30997
1966-70	66697	34368	32329
1971-75	61393	31487	29906
1976-80	51744	26619	25125
1981-85	50660	26030	24629
1986-90	56862	29154	27708
1991-95	60196	30993	29202
1996-00	59522	30598	29043
1996	60927	31490	29437
1997	59801	30724	29077
1998	58352	29870	28482
1999	59298	30468	29421
2000	59234	30436	28798
2001	56696	29041	27655
2002	55434	28325	27109

(Modifisert fra www.ssb.no/aarbok/)

Den relative frekvensen av jentefødsler varierer lite fra år til år og fra femårsperiode til femårsperiode

I perioden 1951-2000 ble det født 2983800 barn i Norge. Av disse var 1449400 jenter

Relativ frekvens av jenter i hele perioden er 48.6%

Den relative frekvensen av seksere er omtrent 1/6 når vi kaster en terning mange ganger

Den relative frekvensen av jenter blant alle nyfødte er omtrent 48.6% hvert år

Det er to eksempler på et fenomen som observeres gang på gang – et fenomen som er en forutsetning for å forstå hva sannsynlighet er:

Vi er interessert i en begivenhet *A* i et stokastisk forsøk. Forsøket gjentas under like betingelser. Da vil den relative frekvensen av *A* nærme seg en grenseverdi når forsøket gjentas mange ganger. Denne grenseverdien er sannsynligheten *P*(*A*)

Sannsynlighet er relativ frekvens i det lange løp

Merk at det er en forutsetning at forsøket gjentas under *like betingelser*

Et eksempel hvor dette ikke er tilfelle er trillingfødsler:

År Årsgj.sn	Levendefødte I alt	Gutter	Jenter	Flerfødsler I alt	Tvilling	Trilling
1951-55	62478	32182	30296	796	787	9
1956-60	63021	32374	30647	738	731	7
1961-65	63989	32992	30997	708	700	8
1966-70	66697	34368	32329	670	663	7
1971-75	61393	31487	29906	572	568	4
1976-80	51744	26619	25125	498	494	4
1981-85	50660	26030	24629	503	495	8
1986-90	56862	29154	27708	653	634	19
1991-95	60196	30993	29202	845	821	24
1996-00	59522	30598	29043	981	957	24
1996	60927	31490	29437	913	886	27
1997	59801	30724	29077	960	931	29
1998	58352	29870	28482	928	903	25
1999	59298	30468	29421	1053	1031	22
2000	59234	30436	28798	1052	1033	16
2001	56696	29041	27655	1035	1016	19
2002	55434	28325	27109	1059	1044	15

Andelen trillingfødsler økte fra midten av 1980 årene på grunn av kunstig befruktning

Sannsynlighet forstått som relativ frekvens i det lange løp kalles enkelte ganger *frekventistisk sannsynlighet*

Merk at denne forståelsen av sannsynlighet forutsetter at forsøket kan gjentas flere ganger

Noen ganger brukes sannsynlighetsbegrepet i situasjoner hvor forsøket ikke kan gjentas.

"Det er 70% sannsynlig at Rosenborg vil slå Stabæk neste søndag"

Sannsynlighet som uttrykker en personlig vurdering kalles *subjektiv sannsynlighet*