**Excel-øvelser i**

**sannsynlighetsregning**

**av**

**Peer Andersen**

**© Peer Andersen 2010**

# Sannsynlighetsregning

Excel er et dataprogram som er utmerket å bruke til simuleringer i sannsynlighetsregning. Vi skal i disse øvelsene (med unntak av øvelse 7) simulere opptil 5000 kast med terninger eller pengestykker. Vi skal også se på hvordan simuleringene samsvarer med de teoretiske modellene. Knut Ole Lysø presenterer i boken *Sannsynlighetsregning – en fagdidaktisk innføring* utgitt på Caspar forlag, flere interessante spill og aktiviteter basert på terningkast og myntkast. Flere av øvelsene i dette kapittelet kan med fordel kombineres med egenaktivitetene som Lysø beskriver i sin bok.

Det er noen grunnleggende formler som går igjen i de aller fleste øvelsene, og som jeg velger å beskrive i innledningen. Det går på hvordan vi kan generere tilfeldige tall og bokstaver, samt hvordan vi kan summere opp f. eks. antall 6'ere. Vi skal også se på hvordan vi kan lage tellevariable.

Generering av tilfeldige tall

I de fleste av disse øvelsene vil dere ha behov for å generere tilfeldige heltall. For eksempel med kast av terning vil det være behov for å generere tilfeldige heltall mellom 1 og 6. Dette kan løses ved å bruke funksjonen TILFELDIG(). Denne funksjonen genererer et tilfeldig tall mellom 0 og 1. Hvis vi multipliserer den med 6, vil vi få et tilfeldig tall mellom 0 og 6. Neste skritt blir å legge til 0,5 slik at vi får et tilfeldig tall mellom 0,5 og 6,5.

Så langt gir funksjonen TILFELDIG()\*6+0,5 et tilfeldig tall mellom 0,5 og 6,5. Vi skal nå avrunde dette tallet til 0 desimaler. Til det kan vi bruke en funksjon som heter AVRUND(*tall*;0). Denne funksjonen avrunder tallet som står der hvor *tall* står til antall desimaler etter semikolonet, i dette tilfellet 0 desimaler. Setter vi sammen disse to funksjonene, får vi funksjonen

=AVRUND(TILFELDIG()\*6+0,5;0)

Dette er en funksjon som genererer et tilfeldig heltall mellom 1 og 6. Vi kan også bruke denne funksjonen til å generere andre heltall. Dersom vi ønsker å generere et heltall mellom 1 og 3, erstatter vi 6 med 3 i formelen over. I enkelte versjoner av Excel finnes det en funksjon som er vesentlig enklere enn den som er beskrevet over. Det er en funksjon som heter

TILFELDIGMELLOM(*tall1;tall2*)

Denne funksjonen gir et heltall mellom *tall1* og *tall2.* For å simulere terningkast kan vi da bruke funksjonen

=TILFELDIGMELLOM(1;6)

Denne funksjonen finnes som sagt bare ikke i alle Excel-versjonene, så den videre fremstilling og filene på den vedlagte CD-ROM’en er basert på den første av formlene.

Generering av tilfeldige bokstaver (f. eks M eller K)

I en av oppgavene vil det være behov for å generere en tilfeldig bokstav, f. eks. M for mynt eller K for krone. Dette kan gjøres ved å kombinere to funksjoner, den som er beskrevet i foregående avsnitt samt en funksjon som heter HVIS. HVIS-funksjonen er bygget opp slik:

HVIS(logisk test;verdi hvis sann;verdi hvis usann)

La oss illustrere bruken med et eksempel. Vi ser på funksjonen

=HVIS(A1=1;”M”;”K”).

Denne funksjonen tester innholdet i rute A1 og sjekker om den er lik 1 eller ikke. Dersom denne er lik 1, får cellen verdien M. Hvis den ikke er lik 1, får den verdien K. Fra forrige avsnitt fant vi ut at funksjonen

=AVRUND(TILFELDIG()\*2+0,5;0)

gir et heltall mellom 1 og 2. (Vi ganger bare med 2 istedenfor med 6). Ved å kombinere disse to funksjonene kan vi få ut verdien M eller K i en celle. Funksjonen som gir dette er:

=HVIS(AVRUND(TILFELDIG()\*2+0,5;0)=1;"M";"K")

Hvis du bruker TILFELDIGMELLOM funksjonen blir funksjonen som generer M eller K slik:

=HVIS(TILFELDIGMELLOM(1;2)=1;"M";"K")

Poenget her er at maskinen først genererer et heltall som er enten 1 eller 2 ved å bruke funksjonen for å generere tilfeldige heltall. Deretter tester maskinen om denne har verdien 1 eller ikke. Dersom den har verdien 1, skrives M i cellen, og hvis ikke (dvs. hvis verdien er 2), så skrives K i cellen.

Summering av f. eks antall 6’ere

Dere vil også få behov for å summere f. eks antall 6’ere. Til det finnes det også en funksjon som er grei å bruke. Det er funksjonen ANTALL.HVIS som vi også brukte på statistikkøvelsene. Da denne funksjonen er så pass sentral, tar vi en liten oppfriskning av den i dette avsnittet. For å illustrere bruken av denne funksjonen, skal vi se på et lite eksempel. Vi skal ta utgangspunkt i regnearket under. Dette arket kaster en terning 20 ganger.



I rute B2 har vi skrevet inn formelen som vi nevnte i forrige avsnitt. Den gir et tilfeldig heltall mellom 1 og 6. Som vi ser her, er det behov for en funksjon som kan summere antall 1’ere, 2’ere osv. Funksjonen som hjelper oss med det er:

=ANTALL.HVIS(B2:B21;"=1")

Det er denne funksjonen som er skrevet inn i rute F2. Den søker gjennom kolonnen fra B2 til B21 og teller hvor mange 1’ere den finner i denne kolonnen. Tilsvarende kan vi gjøre med 2’ere, 3’ere osv. Vi erstatter da bare "=1" med "=2" osv. Vi har skrevet inn "=1". Vanligvis er det imidlertid nok å bare skrive inn tallet 1.

I stedet for å skrive inn funksjonen ”for hånd" kan vi med fordel bruke funksjonsknappen. Klikker vi på denne og leter oss frem til funksjonen ANTALL.HVIS og dobbeltklikker på denne, får vi opp et vindu (se under) der vi kan fylle inn nødvendige data. I området skriver du inn det området vi skal søke gjennom. I vårt tilfelle er det B2:B21. Du kan også merke området i stedet for å skrive det inn. I vilkår skriver du inn hva vi skal søke etter. Når vi skal lete etter 1’ere skriver vi inn 1. Når det er gjort, er det bare å trykke på ok.



Tellevariabler

Hvis du studerer regnearket over, ser du at i kolonne A starter verdiene på 1 og fortsetter oppover. Dersom vi ikke skal gjøre så mange kast, kan vi skrive inn disse tallene hver for seg. Skal vi derimot utføre mange forsøk, f. eks. 1000 kast, så blir det for tungvint å skrive inn alle tallene fra 1 til 1000. Dette kan vi lett unngå ved å bruke en liten formel. Vi tar utgangspunkt i regnearket over og skriver inn tallet 1 i rute A2. I rute A3 skriver vi =A2+1. Denne formelen kan vi så kopiere nedover så langt vi ønsker. Skal vi ha 1000 kast, kopierer vi til 1001, siden vi startet i rute A2. Det enkleste å gjøre når du skal kopiere nedover, er å flytte musen til den ruten som skal kopieres. Flytt musen forsiktig ned til nederste hjørne på høyre side, slik at musepekeren blir et svart lite kryss. Når du har fått frem krysset, holder du venstre musetast nede og drar nedover så langt som du ønsker å kopiere.

Øvelse 1: Kast med 1 terning

Vi skal i denne øvelsen simulere terningkast med 1 terning. Vi skal først bruke Excel til å simulere 100 kast med en terning, så skal vi lage en tabell som viser frekvensen for de ulike utfallene, og til slutt skal vi samle dette i et diagram. Deretter skal vi gjøre det samme med 5000 kast.

Kast med en terning 100 ganger

Vårt mål er å lage et regneark som vist under. Hvis vi ser på kolonnen Kast nr., så skal denne fortsette til 100, selv om bare 24 kast er kommet med på skjermbildet.



Du skal nå lage dette regnearket. Kolonnen Kast nr. er en nummerering av kastene. Den starter på 1 og går til 100. Vi kan lage den ved å skrive inn 1 i ruta A4. I rute A5 skriver vi =A4+1. Denne kopierer vi nedover til vi kommer til kast nummer 100. Kolonnen Terning symboliserer antall øyne på de respektive kastene. Her kan du bruke funksjonen som er beskrevet i innledningen som genererer tilfeldige heltall mellom 1 og 6. Kopier denne nedover til du kommer til kast nummer 100. I kolonnen Antall har vi summert opp antall 1’ere, 2’ere osv. (Husk at regnearket egentlig går til 100 kast.) For å summere opp antall 1’ere osv. kan du bruke ANTALL.HVIS funksjonen som er beskrevet i innledningen. I kolonnen Relativ Frekvens har vi regnet ut den relative frekvensen basert på antall 1’ere, 2’ere osv. Den relative frekvensen finnes ved å dele frekvensen på det totale antall kast. I celle G4 kan du skrive inn =F4/100. Denne kan du kopiere ned til og med celle G9.

Til slutt er alt sammen fremstilt i et stolpediagram. Når du skal lage stolpediagrammet merker du først det du skal fremstille. I denne øvelsen velger vi å fremstille frekvensen. Merk derfor celle E4 til og F9. Klikk deretter på Sett inn og velg Stolpe. Der kan du f. eks velge diagrammet som står helt øverst til venstre. Hvis du vil ha tittel med på diagrammet velger du alternativet som er øverst til venstre i vinduet for Diagramoppsett. Dersom du ikke får dette opp så klikk en eller annen plass i diagrammet og deretter på Diagramverktøy. Diagrammet vi har fått opp er ganske kjedelig med tanke på farger, effekter etc. Vi skal ikke ta noen lang innføring her med hvordan en gjør diagrammet mer fancy, men kort nevne at hvis du høyreklikker på en av stolpene og velger Formater dataserie får du mulighet til å gjøre endringer på stolpene. Hvis du gjøre endringer på bakgrunnen der hvor stolpene står høyreklikker du en eller annen plass i bakgrunnen og velger Formater tegnområdet. Da får du en rekke muligheter til å gjøre det mer fancy.

Det som kan være hensiktsmessig å gjøre er å låse verdiene på y-aksen. Det kan du gjøre ved å høyreklikke på et tall på y-aksen og deretter velge Format akse. Velg Aksealternativer. På Minimum og Maksimum haker du ut for fast og velger henholdsvis 0 og 30.

Kast med en terning 1000 og 3000 ganger

Når vi har konstruert regnearket for 100 kast, er det i grunnen ganske enkelt å konstruere regneark med 5000 kast. Vi følger samme prinsipp som for 100 kast. Når du skal lage regnearket for 5000 kast, er det enkleste å merke hele arket med 100 kast og kopiere det. Gå deretter til ark2 og sørg for at skrivemerket står i rute A1. Klikk på lim inn for å lime inn det du har kopiert fra ark1. Da får du kopiert det som du har gjort over på et nytt ark. Deretter modifiserer du dette arket slik at det gir oss 5000 kast. Du må forlenge kolonnen med Kast nr. slik at denne går til 5000. Det må også gjøres modifisering på ANTALL.HVIS funksjonen slik at den sjekker alle 5000 kastene. Stolpediagrammet i det nye regnearket blir dessverre ikke riktig. Stolpediagrammet som vises er stolpediagrammet for simuleringen med 100 kast. Å modifisere et stolpediagram er ikke alltid så enkelt, og det greieste er faktisk å slette det og så lage et nytt. I forsøket med 5000 kast kan du gjerne lage diagram over den relative frekvensen istedenfor frekvensen slik vi gjorde med 100 kast. Merk da E3 til G9 og klikk på Stolpe. Du får nå fremstilt både frekvensen og den relative frekvensen i samme diagram. Den relative frekvensen er knapt synlig siden den er så liten i forhold til frekvensen. Vi skal fjerne stolpene som viser frekvensen slik at vi kun står igjen med relative frekvensen. Det gjør du ved å høyre klikke på en av stolpene og deretter velge Merk data. Merk deretter Frekvens og klikk på Fjern. Du har nå fått et diagram hvor bare den relative frekvensen er fremstilt. Du kan nå modifisere diagrammet med farger og effekter slik du ønsker.

Oppgave

a) Hvis du skulle kastet en fysisk terning 100 ganger, hvilket av utfallene tror du ville inntruffet flest ganger? Vil det være en favoritt eller vil det være tilfeldig hva som inntreffer flest ganger?

b) Når en trykker på F9-knappen utfører datamaskinen simuleringen på nytt. Det vil si den kaster på nytt igjen og summerer opp de ulike utfallene og viser dette i diagrammet. Gjør dette noen ganger på begge regnearkene du har konstruert. Hva observerer du? Prøv å gi en god forklaring på det som du observerer.

c) Drøft hvordan en slik oppgave kan brukes i grunnskolen og hvordan en slik aktivitet passer inn i forhold til kompetansemålene etter 10. klasse i Kunnskapsløftet.

## Øvelse 2: Kast med 2 terninger

Vi skal i denne øvelsen simulere terningkast med 2 terninger. Vi skal først simulere 100 kast med to terninger og la Excel beregne summen av øynene i hvert enkelt tilfelle. Vi skal deretter lage en tabell som viser frekvensen for de ulike utfallene, og til slutt skal vi samle dette i et diagram. Deretter skal vi gjøre det samme med 5000 kast.

Kast med to terninger 100 ganger

Vårt mål er å lage et regneark som vist under. Hvis vi ser på kolonnen Kast nr., så skal denne fortsette til 100, selv om bare 24 kast er kommet med på skjermbildet.



Oppgaven din blir nå å lage dette regnearket. Kolonnen Terning 1 og Terning 2 symboliserer antall øyne på kastene for terning 1 og terning 2. Her kan du bruke funksjonen som er beskrevet i innledningen som genererer tilfeldige heltall mellom 1 og 6. I kolonnen Sum har vi summert antall øyne på terning 1 og terning 2. I kolonnen Frekvens har vi summert opp hvor ofte vi har fått sum på 2, sum på 3 osv. (Husk at regnearket egentlig går til 100 kast.) For å summere opp hvor ofte vi får en sum på 2, sum på 3 osv., kan du bruke ANTALL.HVIS-funksjonen som er beskrevet i innledningen. Det er imidlertid en del arbeid siden det er 11 ulike utfall, så det kan lønne seg å bruke kopieringsfunksjonen. Da må en imidlertid være litt nøye med hva en gjør, ellers blir det fort feil. Formelen vi skriver i rute G5 er

=ANTALL.HVIS(D4:D103;2).

Kopieres denne, vil du se at i rute G6 har D4 blitt til D5 og D103 har blitt til D104. Det er ikke ønskelig. En kan motvirke dette ved å sette inn dollartegn i cellereferansen. Da vil ikke cellereferansene endre seg ved kopiering. Vi vil derfor bruke formelen

=ANTALL.HVIS($D$4:$D$103;2)

i rute G5. Denne kan vi så kopiere nedover. Vi må passe på at vi etter kopiering endrer 2 til 3 osv. i formelen.

I kolonnen Relativ Frekvens har vi regnet ut den relative frekvensen, på tilsvarende måte som vi gjorde for øvelse 1. Til slutt er alt sammen fremstilt i et stolpediagram. Stolpediagram kan du lage ved å følge samme mal som er beskrevet i øvelse 1.

Kast med to terninger 5000 ganger

Når vi har konstruert regnearket for 100 kast, er det i ganske enkelt å konstruere regneark med 5000 kast. Vi kan følge samme prinsipp som i øvelse 1 ved at vi kopierer det vi har gjort i ark1 over i ark2. Vi modifiserer regnearket på tilsvarende måte som i denne øvelsen. Vi så i sted at vi måtte gjøre en justering på ANTALL.HVIS funksjonen slik at den sjekket alle 5000 kastene. Vi skal her se at dette kan gjøres på en enkel og elegant måte. Merk feltene fra G4 til G14, dvs. de feltene hvor du bruker ANTALL.HVIS funksjonen for å summere opp hyppigheten av de ulike utfallene. Klikk på knappen Søk etter og merk. Det er knappen som ser ut som en kikkert. Velg der Erstatt. Du får nå opp følgende vindu:



I feltet etter Søk etter: skriver du 103, og i feltet Erstatt med: skriver du 5003. Excel vil da endre cellereferansen fra D104 til D5003. Som i foregående øvelse er det enklere å slette diagrammet og lage et nytt enn å prøve å modifisere det som du har fått opp. Lag også her diagram over den relative frekvensen slik som i øvelse 1.

Oppgave

a) Tenk deg nå at du skal kaste to terninger 100 ganger, og registrere hva summen blir hver av gangene. På hvert av kastene kan du få sum fra 2 til 12. Hvilken av summene tror du inntrer oftest? Begrunn svaret.

b) Ved ar du trykker på F9 knappen utfører datamaskinen simuleringen på nytt. Det vil si den kaster på nytt igjen og summerer opp de ulike utfallene og viser dette i diagrammet. Gjør dette noen ganger på begge regnearkene du har konstruert. Hva observerer du? Prøv å gi en god forklaring på det som du observerer.

c) Regn ut hva den teoretiske sannsynligheten blir for få en sum på 2, en sum på 3 og så videre opp til en sum på 12. Hvordan stemmer dette med resultatet av simuleringen din?

## Øvelse 3: Kast med to mynter

Vi skal i denne øvelsen simulere kast med to mynter. Vi skal først ”kaste” to mynter 100 ganger, så lage en tabell som viser frekvensen på de ulike utfallene og til slutt samle dette i et stolpediagram. Deretter skal vi gjøre det samme med 1000 kast.

Kast med to mynter 100 ganger

Vårt mål er å lage et regneark som vist under. Hvis vi ser på kolonnen Kast nr., så skal denne fortsette til 100, selv om bare 24 kast er kommet med på skjermbildet.



Oppgaven din blir nå å lage dette regnearket. Kolonnen Mynt 1 og Mynt 2 symboliserer om vi fikk krone eller mynt på de respektive pengestykkene som vi kastet. I innledningen er det beskrevet en funksjon som genererer tilfeldige bokstav som enten er M eller K. Denne funksjonen kan du bruke. Vi har også med en kolonne som vi har kalt Antall krone. Ved å bruke ANTALL.HVIS funksjonen kan vi få summert opp antall ganger vi får krone på de to pengestykkene. I rute D5 vil funksjonen

=ANTALL.HVIS(B5:C5;"K")

gi det ønskede resultat. Du kan selvsagt også bruke funksjonsveiviseren istedenfor å skrive inn funksjonen.

I kolonnen Frekvens har vi summert opp antall krone-krone, krone-mynt og mynt-mynt. (Husk at regnearket egentlig går til 100 kast.) For å summere opp antall krone-krone osv. kan du bruke ANTALL.HVIS funksjonen som er beskrevet i innledningen. I kolonnen Relativ frekvens, har vi regnet ut den relative frekvensen basert på antall krone-krone osv. Til slutt skal du fremstille dataene i et stolpediagram. Her kan du bruke instruksjonene som er beskrevet i øvelse 1. I forsøket med 100 kast kan du fremstille frekvensen i diagrammet.

Kast med to mynter 5000 ganger

Når vi har konstruert regnearket for 100 kast, er det i grunnen ganske greit å konstruere regneark med 5000 kast. Vi følger samme prinsippet som for 100 kast. Når du skal lage regnearket for 5000 kast, er det enkleste å merke hele arket med 100 kast, kopier det og deretter gå til ark2. Klikk på Lim inn. Da får du kopiert det som du har gjort over på et nytt ark. Deretter modifiserer du dette arket slik at det gir oss 5000 kast. Du må forlenge kolonnen med Mynt 1 og Mynt 2 slik at denne gir oss 5000 kast. Det må også gjøres modifisering på ANTALL.HVIS funksjonen slik at den sjekker alle 5000 kastene. Også her er det greiest å lage nytt diagram fremfor å modifisere det du har kopiert. Lag gjerne et diagram over den relative frekvensen.

Oppgave

a) Tenk deg at du kaster to mynter 100 ganger. Hvilket utfall tror du fremkommer oftest, to krone, to mynt eller en av hver?

b) Ved at man trykker på F9-knappen, utfører datamaskinen simuleringen på nytt. Det vil si den kaster på nytt igjen og summerer opp de ulike utfallene og viser dette i diagrammet. Gjør dette noen ganger på arket med både 100 kast og 5000 kast. Hva observerer du? Er det noe vesentlige forskjeller på 100 kast og 5000 kast? Hva kan eventuelt årsakene være til dette?

c) Regn nå ut eksakt hva sjansen er for å få henholdsvis to krone, to mynt og en av hver. Hvordan stemmer resultatene med simuleringen og det du fant ut i spørsmål b)?

## Øvelse 4: Undersøkelse av den relative frekvens

Vi skal i denne oppgaven undersøke den relative frekvens. Oppgaven er todelt, først en del der vi fysisk skal bruke en terning og deretter en del der vi skal simulere dette i Excel. Den første delen av denne øvelsen er hentet fra *Sannsynlighetsregning – en fagdidaktisk innføring* av Knut Ole Lysø, Caspar forlag.

#### Fysisk kast med terning

Les hele teksten før dere går i gang med arbeidet.

I denne øvelsen skal dere kaste en terning minst 50 ganger. I hvert kast skal dere registrere om kastet ender med en 6’er eller ikke 6’er. Ideen er at dere skal regne ut andelen 6’ere etter hvert kast, sette av denne andelen i et koordinatsystem og finne ut hva som skjer med denne andelen etter hvert som antall kast øker.

I prinsippet kan dette medføre minst 50 utregninger av andelen 6’ere (den relative frekvens). En kan imidlertid forenkle dette arbeidet noe. I eksempelet vist i tabellen under, har en fått 6’er i første, femte og tiende kast.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Kast nummer | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Sekser | x |  |  |  | x |  |  |  |  | x |
| Relativ frekvens | 1,00 |  |  | 0,25 | 0,40 |  |  |  | 0,22 | 0,30 |

I denne tabellen har vi nøyd oss med å regne ut den relative frekvens kun i de tilfeller vi har fått en 6’er og for kastet foran. Grafisk kan vi merke av disse punktene i koordinatsystemet. Da vil den relative frekvens synke fra kast 1 til og med kast 4, stige fra kast 4 til kast 5 for så å synke igjen til og med kast 9 osv. Trekker en forbindelsen mellom disse punktene, vil en på denne måten få en god skisse av forløpet til den relative frekvens.

Fordel oppgavene mellom dere og kast terningen minst 50 ganger slik det er skissert foran.

a) Gi en beskrivelse av forløpet til den relative frekvens på grunnlag av den grafiske fremstillingen dere får fram.

b) Hvordan vil dere forklare det dere finner ut?

c) Hvis dere kaster en terning, hva er sannsynligheten for at dere får en 6’er? Hva er begrunnelsen for svaret? Er det noen sammenheng mellom dette svaret og det dere erfarer med den relative frekvens?

d) Hva kalles denne måten å finne sannsynligheter på?

#### Excelsimulering

Vi skal nå gjøre en simulering av dette i Excel. Vi skal først ta for oss 100 kast og deretter 1000 kast.

###### Simulering av 100 terningkast

Vårt mål er å lage et regneark som vist under. Hvis vi ser på kolonnen Kast nr., så skal denne fortsette til 100, selv om bare 21 kast er kommet med på skjermbildet.



I kolonnen Kast nr. nummerer du kastene fra 1 til 100 fra rute A7 til A106. I rute A6 skiver du 0. Dette er et lite triks vi må gjøre for å få et pent diagram. I kolonnen Verdi er utfallet på hvert enkelt kast gitt. Vi skriver ikke noe i linjen hvor kast nummer 0 står. Du kan bruke funksjonen for å generere tilfeldige tall mellom 1 og 6 som ble beskrevet i innledningen. I kolonnen Rel. frek. har vi fortløpende beregnet den relativ frekvensen for kastserien. For å illustrere hvordan vi kan gjøre dette tar vi en titt på rute C15. I den har vi formelen

=ANTALL.HVIS($B$7:B15;"=6")/A15

Denne formelen går gjennom cellene fra B7 til B15 og teller opp hvor mange 6’ere vi har så langt. Deretter deler den dette antallet på 9 (verdien i A15) for å finne den relative frekvens. Vi er imidlertid interessert i å finne en formel som lett lar seg kopiere, og måten vi kan gjøre det på er at vi i rute C7 bruker formelen

=ANTALL.HVIS($B$7:B7;"=6")/A7.

Det som skjer når vi kopierer denne formelen nedover, er at $B$7 beholdes og at B7 og A7 endres etter hvert som vi kopierer. Når vi kopierer denne formelen til C8, vil vi få funksjonen ANTALL.HVIS($B$7:B8;"=6")/A8. Tilsvarende vil vi i rute C15 få formelen over når vi fortsetter å kopiere.

Vi har også tatt med siste verdien, dvs. den relative frekvensen vi har etter 100 kast og plassert denne like under diagrammet. Det er gjort for å få med sluttresultatet sammen med diagrammet. Dette kan gjøres ved å skrive =C106 i en av cellene like under diagrammet.

Til slutt skal vi fremstille dette i et linjediagram. I diagrammet er vi interessert i at Kast nr. skal være x-aksen og Rel. frek. skal være y-aksen. Kolonnen med verdi skal vi ikke bruke i diagrammet.

Start med å merke cellene C6 til C106. Klikk deretter på Sett inn og velge Punkt. Velg der diagrammet som heter Punktdiagram med utjevnende linjer. Det er i 2 linjen til venstre. Du skal nå ha fått opp et linjediagram. Det kan være hensiktsmessig å låse verdiene på y-aksen også her. Klikk på et av tallene på y-aksen og velg formater akse. Der velger du 0 på minimum og f. eks 0,5 på maksimum.

5000 kast

Når vi har konstruert regnearket for 100 kast, er det i grunnen ganske greit å konstruere regneark med 1000 kast. Vi følger samme prinsippet som for 100 kast. Når du skal lage regnearket for 1000 kast, er det enkleste å merke hele arket med 100 kast, kopiere det og deretter gå til ark2. Klikk på lim inn. Da får du kopiert det som du har gjort over på et nytt ark. Deretter modifiserer du dette arket slik at det gir oss 1000 kast. Du må forlenge kolonnen med kast nr. slik at denne går til 1000.

Også her er det nok enklest å lage et nytt diagram fremfor å begynne å tilpasse det gamle. Det er lett for at det da blir feil.

Oppgave

a) Gjør noen simuleringer med 100 kast. Hvordan stemmer dette overens med det som du gjorde fysisk med terningen?

b) Gjør noen simuleringer både med 100 kast og 1000 kast. Hva observerer du? Prøv å forklare det som du observerer.

c) Hvis du kaster en terning, hva er sannsynligheten for at du får en 6’er? Hvordan samsvarer dette med simuleringene du har gjort?

## Øvelse 5: Mercedes – geit-problemet

Mercedes–geit problemet er en klassisk problemstilling som har skapt mye diskusjon opp gjennom årene. Så vel eksperter som vanlige folk har diskutert problemstillingen flittig. Problemstillingen tar utgangspunkt i et amerikansk fjernsynsshow der en deltaker kommer inn i et rom med tre dører. Bak to av dørene befinner det seg en geit og bak den tredje døren skjuler det seg en Mercedes. Programlederen ber deltakeren om å velge en av dørene, og vedkommende får i utgangspunktet det som er bak døren.

a) Hva er sannsynligheten for at den døren deltakeren velger skjuler Mercedesen? Hva er sannsynligheten for at det er en geit bak den valgte døren?

Programlederen vet hva som skjuler seg bak dørene, og når deltakeren har valgt en dør (men før den blir åpnet), inngår det i showet at programlederen åpner en av de andre dørene. Denne viser alltid en geit. Deltakeren får nå velge blant to alternativer. Enten kan han åpne den døren han har valgt eller han kan bytte dør.

b) Dersom du er deltaker i showet, hva ville du ha gjort? Ville du byttet dør, eller ville du åpnet den du hadde valgt i utgangspunktet? Begrunn svaret.

c) Anta nå at du i utgangspunktet hadde valgt en dør med en geit bak. Hva er sjansen for at du vinner en Mercedes hvis du bytter dør?

d) Vi antar nå at du først valgte døren med Mercedesen bak. Hva er sjansen for at du vinner Mercedesen hvis du bytter dør?

e) Hva slags konklusjon vil du trekke? Bør deltakerne i programmet bytte dør, eller bør de velge den døren de i utgangspunktet har valgt?

Excel simulering

Vi skal nå gjøre en simulering i Excel av dette problemet. Vi skal simulere 1000 forskjellige show og se hvor mange ganger vi får Mercedes ved å bytte dør, og hvor mange ganger vi får Mercedes ved å stå i ro.

Denne simuleringen er noe mer krevende å lage enn de vi hittil har jobbet med, men den skal likevel være mulig å få til. Vi skal se stegvis på hvordan vi kan konstruere regnearket. På neste side har vi kopiert inn et regneark som viser hvordan regnearket vi skal lage kan se ut.



Vi har her bare tatt med de 21 første showene, men regnearket består egentlig av 1000 show.

Vi starter med å skrive inn teksten som skal inn, samtidig som vi også gir regnearket et nytt navn f. eks. Mercedes–Geit. Kolonnen Show er bare en tellekolonne som nummerer de 1000 showene vi skal simulere. Vi skal nå gå i gang med å bestemme hvilke dører som skjuler geitene og hvilken dør som skjuler Mercedesen. For å få dette til, må vi lage en liten hjelpekolonne. Den skal vi lage i ark2 slik at den ikke blir synlig i hovedarket vårt. Vi skal bruke cellene fra A7 til A1006 i ark2 til dette. Vi skal i disse cellene generere et heltall mellom 1 og 3. Dette gjøres ved å bruke funksjonen for generering av tilfeldige tall som er beskrevet i innledningen. Dersom dette tallet er 1, betyr det at Mercedesen befinner seg bak dør 1 og at det er geit bak dør 2 og 3. Dersom dette tallet er 2 er Mercedesen bak dør 2, og tilsvarende om tallet er 3, så er Mercedesen bak dør 3.

På grunnlag av dette har vi i kolonne Dør 1, Dør 2 og Dør 3 skrevet ned hva som befinner seg bak de respektive dører. Til dette bruker vi HVIS funksjonen. I rute B7 er formelen

=HVIS(ark2!A7=1;"M";"G")

skrevet inn. Excel-programmet sjekker da hva innholdet er i rute A7 på ark2. Dersom dette er 1, så skriver maskinen M i rute B7. Dersom A7 er 2 eller 3, vil maskinen skrive inn en G istedenfor. Helt tilsvarende formel skriver vi i rute C7 og D7, bare at der sjekker vi om A7 er henholdsvis 2 eller 3.

Kolonnen Valgt dør symboliserer hvilken dør deltakeren har valgt. Dette gjøres ved å bruke funksjonen for å generere et tilfeldig heltall mellom 1 og 3. Kolonnen Bytte viser hva deltakeren får i gevinst dersom han velger å bytte dør. Dette kan beregnes enkelt ved å bruke HVIS funksjonen. I rute G7 bruker vi formelen

=HVIS(E7=ark2!A7;"G";"M")

Det Excel her gjør er at programmet sjekker den døren vi har valgt, som er E7, mot hvilken dør Mercedesen er gjemt bak. Dersom vi har valgt samme dør som Mercedesen er gjemt bak (A7 på ark2 er lik E7), vil vi ende opp med en geit i gevinst. Dersom vi har valgt en dør der Mercedesen ikke er (A7 på ark2 er ikke lik E7), så vinner vi Mercedesen.

I kolonnen Stå skal Excel skrive inn gevinsten vi får dersom vi velger å stå i ro. Dette vil alltid være det motsatte av den gevinsten vi får dersom vi bytter dør, og vi kan derfor bruke HVIS-funksjonen også her. Funksjonen

=HVIS(G7="M";"G";"M")

gir oss gevinsten dersom vi velger å ikke bytte dør.

I feltet til høyre på regnearket har vi beregnet hvor mange ganger vi vil vinne Mercedes i løpet av de 1000 showene ved å bytte dør og ved å stå i ro. Tilsvarende har vi også beregnet hvor mange geiter vi vil vinne dersom vi bytter dør og dersom vi står i ro. Vi kan bruke ANTALL.HVIS-funksjonen til beregne dette. Til slutt har vi beregnet den relative frekvensen

Merknad: Vi kunne strengt tatt droppet kolonnene Dør 1, Dør 2 og Dør 3. De har ingen funksjon i forhold til beregningene som gjøres. (Hvorfor?). Vi tar dem likevel med for å bedre oversikten i programmet. Uten disse kolonnene blir det fort svært vanskelig å forstå hvordan simuleringen egentlig fungerer.

Oppgave

a) Vi har nå simulert 1000 show. Dersom vi raskt ønsker å simulere 1000 nye show, trykker du på F9-tasten. Gjør dette noen ganger. Dersom du hadde deltatt i showet, ville du byttet dør, eller ville du holdt fast ved den du først valgte?

b) Hvordan samsvarer simuleringen med dine betraktninger i første del av oppgaven? Kan en slik simulering være egnet til å overbevise tvilere om at det er lurt å bytte dør?

## Øvelse 6: Kast med 5 terninger

Vi skal i denne øvelsen simulere terningkast med 5 terninger der vi i hvert enkelt tilfelle skal se på hvor mange av terningene som gir 6’ere. Vi skal først simulere 100 kast med 5 terninger, og så skal vi lage en tabell som viser frekvensen på de ulike utfallene. Til slutt skal vi samle dette i et diagram. Deretter skal vi gjøre det samme med 5000 kast.

Kast med 5 terninger 100 ganger

Vårt mål er å lage et regneark som vist under.



Hvis vi ser på kolonnen Kast nr., så skal denne fortsette til 100, selv om vi på skjermbilde bare har fått med 23 kast. Oppgaven blir nå å lage dette regnearket. Kolonnene T1 til T5 symboliserer verdien på de enkelte terningene. Her kan du bruke funksjonen som er beskrevet i innledningen der vi kan generere tilfeldige heltall mellom 1 og 6. I kolonnen Antall 6’ere har vi summert opp hvor mange av de 5 terningene som gir 6’ere. Dette kan gjøres ved å bruke ANTALL.HVIS funksjonen. I kolonnen Frekvens har vi summert opp hvor mange av kastene som gir null 6’ere, en 6’er, to 6’ere osv. Dette gjøres ved å bruke ANTALL.HVIS funksjonen. (Husk her at regnearket egentlig går til 100 kast.) Til slutt er alt sammen fremstilt i et stolpediagram. Stolpediagrammet lager vi etter samme mønster som for øvelse 1. Imidlertid kan det tenkes at du får opp et diagram som vist på neste side

Problemet her er at vi har fått to kolonner og at verdiene på x-aksen ikke stemmer. Grunnen til at dette har skjedd er at den tror at kolonnen med Antall 6’ere også skal fremstilles som stolper. Excel skjønner ikke at dette er verdiene som skal brukes på x-aksen. Vi kan imidlertid fikse på dette ved å høyreklikke på en av stolpene. Velg der Merk data. Klikk på Antall 6’ere og deretter på Fjern. Klikk så på Rediger og merk det som skal være x-aksen vår, det vil si celle I5 til I10. Klikk deretter på ok. Du kan nå eventuelt modifisere diagrammet med farger og effekter.

Kast med 5 terninger 5000

Når vi har konstruert regnearket for 100 kast, så er det i grunnen ganske greit å konstruere regneark med 5000 kast. Vi følger samme prinsippet som for 100 kast. Når du skal lage regnearket for 5000 kast, er det enkleste å merke hele arket med 100 kast, kopiere det og deretter gå til ark2. Klikk på Lim inn. Da får du kopiert det som du har gjort over på et nytt ark. Deretter modifiserer du dette arket slik at det gir oss 5000 kast. Du må forlenge kolonnen med Kast nr. slik at denne går til 5000. Det må også gjøres modifisering på ANTALL.HVIS funksjonen slik at den sjekker alle 5000 kastene. Som i tidligere øvelser så er det enklest å slette diagrammet og lage et nytt ett fremfor å prøve å modifisere det gamle. Lag gjerne diagram for relativ frekvens også her.

Oppgave

a) Når du triller 5 terninger kan du få fra 0 til 5 seksere. Hva tror du det er størst sjanse for å få?

b) Når du trykker på F9-knappen utfører datamaskinen simuleringen på nytt. Det vil si den kaster på nytt igjen og summerer opp de ulike utfallene og viser dette i diagrammet. Gjør dette noen ganger på begge regnearkene du har konstruert. Hva observerer du? Prøv å gi en god forklaring på det som du observerer.

c) Regn ut hva sjansen er for å få henholdsvis 0, 1, 2, 3, 4 og 5 seksere. Hva observerer du? Hvordan stemmer dette med forsøkene du har gjort i spørsmål b)?

d) Å få yatzy på første kast er noe som ikke skjer så ofte. Kan du bruke programmet til å gi et anslag for hva sannsynligheten er for å få yatzy i seksere på første kast? Gi også et anslag for sjansen å få yatzy på første kast.

## Øvelse 7: Binomisk og hypergeometrisk fordeling

Denne oppgaven handler i hovedsak om at vi skal gjøre sammenlikninger mellom den hypergeometriske fordelingen og den binomiske fordelingen ved hjelp av Excel. Vi skal på grunnlag av dette kunne si noe om når den binomiske fordelingen er en god tilnærming til den hypergeometriske modellen. Før vi starter med selve Excel-arbeidet, skal vi starte med en liten arbeidsoppgave der vi ikke skal bruke Excel.

Del 1. Sammenlikning av binomisk og hypergeometrisk fordeling.

Ved Høgskolen i Finnmark er det ca. 2000 studenter. Av disse antar vi at 1200 er kvinner. Nylig har det vært valg til høgskolestyret, og studentene skulle velge 3 representanter. Vi tenker oss at valget skjer ved loddtrekning blant alle studentene.

Vi skal i denne oppgaven se på hvor stor sannsynligheten er for at det blir valgt inn to kvinner og en mann i styret hvis utvelgelsen skjer ved loddtrekning. To studenter som heter Fiskum og Elvestrand er uenige om hvordan dette skal gjøres. Fiskum og Elvestrand argumenterer slik:

Fiskum

Han setter opp følgende uttrykk:

$$\frac{\left(\begin{array}{c}1200\\2\end{array}\right)∙\left(\begin{array}{c}800\\1\end{array}\right)}{\left(\begin{array}{c}2000\\3\end{array}\right)}=0,4323$$

Han argumenterer med at totalt kan vi trekke ut 3 av 2000 studenter på $\left(\begin{array}{c}2000\\3\end{array}\right)$ måter. Av disse kombinasjonene vil det være $\left(\begin{array}{c}1200\\2\end{array}\right)∙\left(\begin{array}{c}800\\1\end{array}\right)$ som inneholder 2 jenter og 1 gutt. Sannsynligheten skulle derfor bli som i uttrykket over.

Elvestrand

Hun setter opp følgende uttrykk:

$$\left(\begin{array}{c}3\\1\end{array}\right)∙0,6∙0,6∙0,4=0,4320$$

Hun argumenterer med at det er 60 % jenter og 40 % gutter på skolen. Av den grunn vil sannsynligheten for at første person som trekkes ut er en jente være 0,6. Sannsynligheten for at andre person er en jente vil også være 0,6 og sannsynligheten for at siste person er en gutt vil da være 0,4. Siden det ikke betyr noe om gutten blir trukket ut som nummer 1, 2 eller 3 så må vi multiplisere uttrykket $0,6∙0,6∙0,4$ med $\left(\begin{array}{c}3\\1\end{array}\right)$. Dermed får vi uttrykket over.

Oppgave

Det er ikke tilfeldig at disse to metodene gir tilnærmet likt svar. Imidlertid er det bare en av dem som er matematisk korrekt. Drøft hvilken av metodene som er matematisk korrekt, og hvorfor de to metodene likevel gir tilnærmet det samme svaret.

Del 2. Undersøkelse ved hjelp av Excel.

Modellene som ble brukt i innledningsoppgaven er hypergeometrisk modell (Fiskum) og binomisk modell (Elvestrand). Vi skal nå bruke Excel til å gjøre en mer systematisk undersøkelse av når det kan være fornuftig å bruke en binomisk modell som tilnærmelse til den hypergeometriske modellen. Vi kan tenke oss en situasjon der vi skal trekke ut 15 fiskekort i Altaelven. Vi skal se på hva sjansen er for å trekke ut henholdsvis 0 kvinner, 1 kvinne, 2 kvinner osv. opp til og med 15 kvinner. Vi forutsetter da at vi vet hvor mange som er med i trekningen og hvor mange av dem som er kvinner. Dette er en hypergeometrisk fordeling, og vi skal nå sammenlikne den med en binomisk fordeling. Vi skal lage regnearket slik at vi oppgir det totale antallet som er med i trekningen samt hvor mange av dem som er kvinner. Regnearket vårt skal se ut omtrent som vist på under.



Her har vi brukt et eksempel med 100 deltakere der 50 av dem er kvinner. Disse verdiene har vi lagt inn i rute D3 og D4. Vi starter med å lage en liten overskrift, samt at vi skriver inn teksten i rute A3 og A4. Vi skal så lage tre kolonner. En tellekolonne som vi har kalt for x, en for den hypergeometriske fordelingen og en for den binomiske fordelingen. Vi lager x-kolonnen slik vi tidligere har gjort, ved å bruke en tellevariabel. Til kolonnen som er merket med hyp skal vi bruke en funksjon som ligger i Excel og som regner ut en hypergeometrisk fordeling. Klikk på -knappen. Velg deretter statistisk og deretter HYPERGEOM.FORDELING. Du vil da få opp følgende dialogboks



Utvalg\_s er antall suksesser i utvalget. Når vi står i rute B7, er vi interessert i å finne ut sjansen for 0 suksesser. Vi kan skrive inn 0 etter Utvalg\_s, men mer elegant vil det være å skrive inn A7. Det gjør at formelen kan kopieres nedover. Utvalgsstørrelse er hvor stort utvalget er. I vårt tilfelle vil det si hvor mange laksekort vi skal trekke ut. Vi skal trekke ut 15 kort så vi skriver inn 15 på utvalgsstørrelse. De to nederste feltene er relatert til hele populasjonen. Suksesser er antall suksesser i hele populasjonen. I vårt tilfelle vil det si antall kvinner som deltar i trekningen. Vi har lagt denne verdien inn i rute D4 så vi skriver $D$4 i feltet for Suksesser. Merk at vi bruker dollartegn, da vi ikke ønsker at verdien skal forandre seg når vi kopierer. Populasjonsstørrelse forteller hvor mange som er med i trekningen totalt. I vårt tilfelle har vi skrevet det inn i rute D3. Vi setter derfor $D$3 i feltet med populasjonsstørrelse. Til slutt klikker du på ok og kopierer formelen nedover.

Vi skal nå regne ut de respektive sannsynligheter ved å bruke en binomisk modell. Vi leter oss frem til funksjonen BINOM.FORDELING ved å bruke samme fremgangsmåte som i sted. Vi får da opp følgende dialogboks.



Antall\_s er antall vellykkede forsøk. I vårt tilfelle vil det si hvor mange kvinner som får fiskekort. Når vi står i rute C7, kan vi skrive inn A7 i dette feltet. Feltet Forsøk viser antall forsøk som gjøres. I vårt tilfelle trekker vi ut 15 fiskekort og vi skriver derfor 15 i dette feltet. Feltet Sannsynlighet\_s sier hvor stor sannsynligheten er for suksess i hvert enkelt forsøk. Denne kjenner vi ikke eksakt. Resultatet av trekningen av første fiskekort vil påvirke utfallet på neste. Vi gjør en tilnærming her ved å si at sannsynligheten er antall kvinner i populasjonen delt på det totale antallet i populasjonen. I vårt tilfelle blir sannsynligheten =$D$4/$D$3, da vi ønsker å bruke verdiene som vi legger inn i starten på arket. På kumulativ skriver du USANN. Når formlene er skrevet inn i rute B7 og C7, kan de kopieres nedover.

Til slutt så skal du fremstille dette i søylediagram. Bruk beskrivelsene som er gitt i de tidligere øvelsene til å lage diagrammet.

Oppgave

Du skal nå teste ut regnearket du har laget. Vi har i starten valgt å la hele populasjonen være 100 og antall kvinner være 50. Gjør nå en mer systematisk undersøkelse der du lar populasjonen variere fra en liten populasjon på 30 og opp til en stor på 10000. La også kvinneandelen variere fra 10 % og opp til 50 %. Hva finner du ut? Når vil den binomiske modellen være en god tilnærming til den hypergeometriske modellen?