Välkomna till Matematisk statistik och diskret matematik!

Andreas Nordvall Lagerås

rum:	L3049
mottagningstid:	tisdagar kl. 9–11, 13–15 (eller enl. överenskom.)
telefon:	772 5337
mejl:	norand@chalmers.se
kurshemsida:	
	www.math.chalmers.se/Stat/Grundutb/CTH/mve050/0708/

- Tre obligatoriska och ganska omfattande grupparbeten ingår i kursen (se hemsidan).
- Se till att också räkna de föreslagna talen (på vecko-pm:et) inför assistenternas genomgångar.

Under kursen kommer vi att gå igenom

- Grundläggande sannolikhetsteori (idag).
- Stokastiska (dvs slumpmässiga) variabler, vars värde beror på något slumpfenomen.
- Stokastiska processer: slumpmässiga skeenden i tiden.
- Statistik: hur man på bästa sätt drar slutsatser från data och hur säker man kan vara på resultaten.
- Diskret matematik, särskilt kombinatorik och så kallade genererande funktioner som är användbart inom sannolikhetsteorin.

Exempel på stokastiska variabler

- Antal prickar nästa tärning jag kastar kommer att visa.
- Börsens upp- eller nedgång i procent idag.
- Nederbördsmängden nästa år.

Exempel på stokastiska processer

- Börskurser.
- Temperaturkurva.
- Datatrafik över internet.

Använd er intuition.

- Sannolikheter ligger mellan 0 och 1, och uttrycks därför ofta i procent.
- Sannolikhet nära 0 betyder att en händelse är osannolik.
- Sannolikhet nära 1 betyder att den är sannolik.
- Sannolikhet lika med 0.5 betyder lika sannolik som osannolik, tänk på en mynsingling (krona/klave).

Fråga er själva, i de fall det är möjligt: ”Är resultatet rimligt?”

$P(A)$ utläses ”sannolikheten för (att) A.”

Begreppet sannolkhet kan tolkas på flera sätt.
Vi kan säga att sannolikheten är *personlig* om olika personer kan tillskriva den olika värden. Särskilt om försöket *inte kan upprepas*. Exempel:

- Kommer den globala medeltemperaturen höjas 4°C? Olika experter kan ge olika uppskattningar av sannolikheten.
- Vadslagning kräver nästan att folk har olika uppfattningar om sannolikheterna för att den ska bli av.

Vi kan ha en *klassisk* tolkning om det är så att alla utfall, dvs alla enskilda saker som kan inträffa, är *lika sannolika*. Då är $P(A) = n(A)/n(S)$, där

- $n(A)$ är antalet utfall som leder till händelsen A,
- $n(S)$ är totala antalet utfall.

Exempel: Tärningskast med vanlig tärning. $A =$ den visar udda.
$n(A) = 3$, eftersom de tre utfallen $1, 3, 5$ är udda.
$n(S) = 6$, eftersom det finns sex möjliga utfall.
Alltså är $P(A) = \frac{3}{6} = \frac{1}{2}$.

Här är ofta kombinatorik användbart.

Vi talar om en *frekventistisk* tolkning av sannolikheter om vi tycker att $P(A) \approx f(A)/n$ där

- $f(A)$ är antalet gånger A inträffar då
- vi *upprepar* ett visst försök n gånger.

. . . men $f(A)$ är ju slumpmässigt i sig. . . ?
Vi menar egentligen

$$\lim_{n\to\infty} \frac{f(A)}{n} = P(A).$$

Vi kan också tala om ett *axiomatiskt* förhållningssätt.

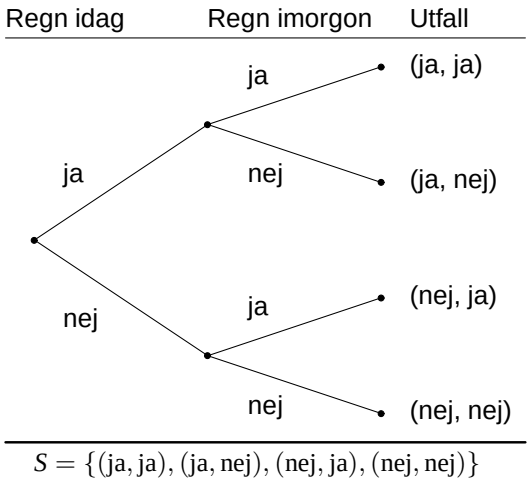
- Ställer bara minimala krav på vad ”sannolikhet” är, dvs hur funktionen $P(A)$ skall bete sig.
- Knyter samman de tidigare nämnda tolkningarna.
- Är i sig tolkningsfri.

Terminologi

- Ett *experiment* leder till något visst *utfall* (även om vi inte kan upprepa experimentet).
- Mängden av alla möjliga utfall kallas för *utfallsrum* (eng. sample space) och betecknas S.
- Olika utfall kan kombineras till *händelser* som alltså är delmängder av S.

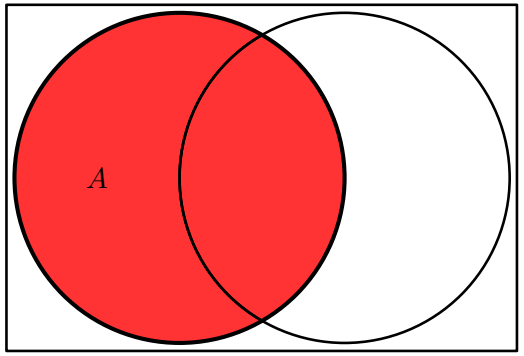
Ex. Tärningskast. Utfallen är $1, 2, 3, 4, 5, 6$. Utfallsrummet är $S = \{1, 2, 3, 4, 5, 6\}$.
Händelser: ”slå mindre än 4” $= \{1, 2, 3\}$, ”slå jämnt” $= \{2, 4, 6\}$, ”slå fyra” $= \{4\}$, etc.

Utfallsrummet kan beskrivas med ett träd om det genereras av flera olika ”val”. Ex: Vädret idag och imorgon.



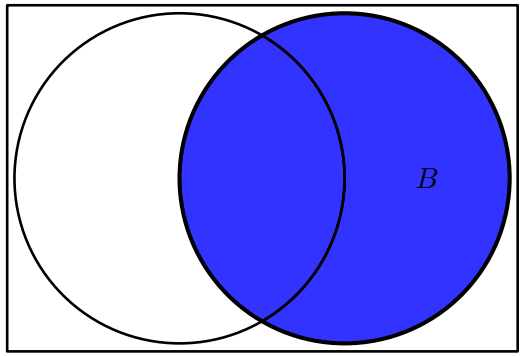
$S = \{(\text{ja}, \text{ja}), (\text{ja}, \text{nej}), (\text{nej}, \text{ja}), (\text{nej}, \text{nej})\}$

Venn-diagram är också väldigt nyttiga för att förstå mängder/händelser.



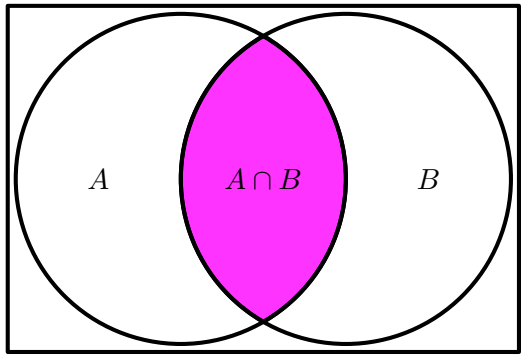
Mängden A

Venn-diagram är också väldigt nyttiga för att förstå mängder/händelser.



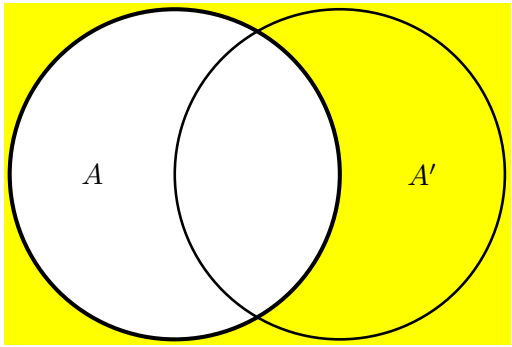
Mängden B

Venn-diagram är också väldigt nyttiga för att förstå mängder/händelser.



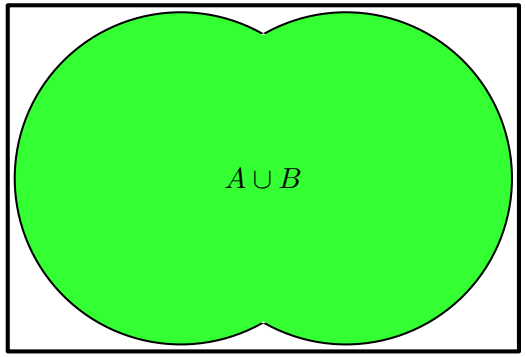
Snittet $A \cap B =$ ”Både A och B”

Venn-diagram är också väldigt nyttiga för att förstå mängder/händelser.



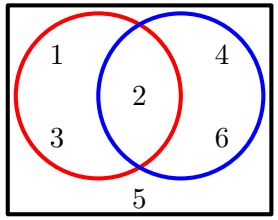
Komplementet: $A' =$ ”Inte A”

Venn-diagram är också väldigt nyttiga för att förstå mängder/händelser.



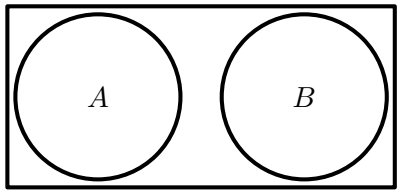
Unionen $A \cup B =$ ”A och/eller B”

Exempel på Venn-diagram för tärningskast.
$A =$ ”slå mindre än 4” $= \{1, 2, 3\}$, $B =$ ”slå jämnt” $= \{2, 4, 6\}$,
$A \cap B =$ ”slå mindre än fyra och jämnt” $= \{2\}$,
$A \cup B =$ ”slå mindre än fyra eller slå jämnt” $= \{1, 2, 3, 4, 6\}$,
$(A \cup B)' =$ ”varken slå mindre än fyra eller slå jämnt” $= \{5\}$.



Figur: Låga i röda cirkeln, jämna i blå.

Två händelser A och B är disjunkta om $A \cap B = \emptyset := \{\}$.
Händelserna $A_1, A_2, \ldots$ är disjunkta om $A_i \cap A_j = \emptyset$ för $i \neq j$.
(Det är fel i bokens definition av ”mutually exclusive events”.)



Figur: Två disjunkta mängder A och B överlappar inte varandra.

Ex. Tärningskast. ”Slå udda” och ”slå 4” är disjunkta händelser.

Ex. (Multinomialkoefficienter) $\binom{n}{n_1,\ldots,n_k}$ = antal sätt man kan dela upp n föremål till k personer så att person i får n_i föremål.

$$\begin{aligned}\binom{n}{n_1,\ldots,n_k} &= \binom{n}{n_1}\binom{n-n_1}{n_2}\cdots\binom{n_k}{n_k} \\ &= \frac{n!}{n_1!(n-n_1)!}\frac{(n-n_1)!}{n_2!(n-n_1-n_2)!}\cdots\frac{n_k!}{n_k!0!} \\ &= \frac{n!}{n_1!\cdots n_k!}\end{aligned}$$

Kombinatorikrepetition

- $n! = n\cdot(n-1)\cdots 2\cdot 1 =$ antal sätt man kan ordna n stycken föremål. ($0! = 1$).
- $\binom{n}{k} = \frac{n!}{k!(n-k)!} =$ antal sätt man kan välja k stycken föremål bland n olika, utan att bry sig om de k styckens inbördes ordning.
- Multiplikationsprincipen: Om man har k omgångar och i omgång nr i har n_i olika val så finns det $n_1 \cdot n_2 \cdots n_k$ olika kombinationer av val man kan göra.

Sannolikhetsaxiomen (Kolmogorov)

All modern sannolikhetsteori bygger på Kolmogorovs sannolikhetsaxiom från 1933.

1. $P(S) = 1$.
2. $P(A) \geq 0, A \subseteq S$.
3. $P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_1) + \cdots$ om $A_1, A_2, \ldots$ alla är disjunkta.

Några enkla och viktiga satser

- $P(A') = 1 - P(A)$.
  Bevis: $S = A \cup A'$ där A och A' är disjunkta $\Rightarrow$
  $1 \overset{1.}{=} P(S) = P(A \cup A') \overset{3.}{=} P(A) + P(A')$. $\square$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
  Bevis: $A \cup B = A \cup (A' \cap B)$, $B = (A \cap B) \cup (A' \cap B)$.

$$\begin{aligned} P(A \cup B) &\overset{3.}{=} P(A) + P(A' \cap B) \\ &= P(A) + P(A' \cap B) + P(A \cap B) - P(A \cap B) \\ &\overset{3.}{=} P(A) + P(B) - P(A \cap B). \quad \square \end{aligned}$$

"$P(A) + P(B)$ räknar $P(A \cap B)$ två gånger om."

Betingning

Om vi vet att en viss händelse har inträffat så kan det påverka sannolikheten för andra händelser.

Ex. Jag kastar en tärning och den ett högt värde: $4, 5$ eller 6.
Vad är sannolikheten att den visar udda? Rimligtvis $\frac{1}{3}$.
Man skriver $P(\text{udda}|\text{högt})$ och säger "sannolikheten för udda *givet* högt".
Obs: $\frac{P(\text{udda} \cap \text{högt})}{P(\text{udda})} = \frac{P(\{5\})}{P(\{4,5,6\})} = \frac{1/6}{3/6} = \frac{1}{3}$.

Allmän definition:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ då } P(B) > 0.$$

Oberoende händelser

En tolkning av betingning är att man tillför information. Men om information om att en viss händelse B har inträffat inte påverkar sannolikheten för A så är/inträffar A och B *oberoende* av varandra.

$$P(A) = P(A|B) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = P(A)P(B).$$

Definition: A och B är oberoende om $P(A \cap B) = P(A)P(B)$.

Ex. Dra ett kort från en välblandad kortlek. Händelserna "ess" och "klöver" är oberoende eftersom
$P(\text{klöver ess}) = \frac{1}{52} = \frac{1}{4} \cdot \frac{1}{13} = P(\text{klöver})P(\text{ess})$.

Ex. Dra ett kort från en välblandad kortlek där spader två saknas. Då är "ess" och "klöver" *inte* oberoende:
$P(\text{klöver ess}) = \frac{1}{51} \neq \frac{52}{51} \cdot \frac{1}{51} = \frac{13}{51} \cdot \frac{4}{51} = P(\text{klöver})P(\text{ess})$.

Det finns några enkla följder av definitionen av betingning.

- Kedjeregeln: $P(A \cap B) = P(A|B)P(B)$.
- Lagen om total sannolikhet:

$$\begin{aligned} P(A) &= P(A \cap B_1) + \cdots + P(A \cap B_n) \\ &= P(A|B_1)P(B_1) + \cdots + P(A|B_n)P(B_n), \end{aligned}$$

  om $B_1, \ldots, B_n$ är disjunkta och $B_1 \cup \cdots \cup B_n = S$.
  Exempel: $B_1 = B, B_2 = B'$.

Den viktigaste följden är *Bayes sats*:

$$\begin{aligned} P(B_i|A) &= \frac{P(A \cap B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{P(A)} \\ &= \frac{P(A|B_i)P(B_i)}{P(A|B_1)P(B_1) + \cdots + P(A|B_n)P(B_n)}. \end{aligned}$$

Vi kan alltså på sätt och vis ”kasta om” betingningen.

Stokastiska variabler

- Stokastiska variabler är variabler som beror på utfallet av ett ”experiment”.
- Ex: Antalet prickar nästa gång jag kastar en tärning.
- Konvention: Stor bokstav ($X, Y, \ldots$) betecknar själva stokastiska variabeln, och liten bokstav betecknar dess utfall ($x, y, \ldots$).

Exempel: Spamfilter

- Bygg upp två listor med alla ord ur alla inkommande mejl, en för spam och en för icke-spam.
- $B =$ ”nytt mejl är spam”. Vi kan skatta $P(B)$ genom proportionen av spam i all tidigare mejl.
- Tag ett ord i det nya mejlet. För varje ord ”A” som finns i listorna kan vi skatta $P(A|B)$ från spam-listan och $P(A|B')$ från icke-spam-listan.
- Genom Bayes sats kan vi få fram $P(B|A)$, den *betingade* sannolikheten för att mejlet är spam givet att vi observerar ordet A.
- Ex: $P(B) = 0.7$, $P(\text{”Dear”}|B) = 0.01$, $P(\text{”Dear”}|B') = 0.001$ ger

$$P(B|\text{”Dear”}) = \frac{0.01 \cdot 0.7}{0.01 \cdot 0.7 + 0.001 \cdot 0.3} \doteq 0.96 > 0.7 = P(B)$$

Diskreta stokastiska variabler

- Definition: En stokastisk variabel är *diskret* om den bara kan anta högst ett uppräkneligt antal värden.
- Motsats: *Kontinuerliga* stokastiska variabler.
- Ex. (Diskret) Antal prickar vid tärningskast (sex utfall).
- Ex. (Diskret) Antal myntsinglingar tills första klaven (uppräckneligt oändligt antal utfall: $1, 2, \ldots$).
- Ex. (Kontinuerlig) Tid (icke-avrundad) tills bussen kommer nästa gång (överuppräkneligt oändligt antal utfall: $\mathbb{R}_+$).

Täthetsfunktion

- Definition: För diskreta stokastiska variabler är täthetsfunktionen (även kallad sannolikhetsfunktionen)

$$f(x) = P(X = x).$$

- Nödvändiga och tillräckliga villkor för att f är en (diskret) täthetsfunktion:
  - $f(x) \geq 0$ (pga axiom 2).
  - $\sum_x f(x) = 1$ (pga axiom 1).

Fördelningsfunktion, F

- Definition:

$$F(x) = P(X \leq x) = P(X = x) + P(X = x-1) + \cdots = \sum_{k \leq x} P(X = k).$$

- Ex. Tärning. $F(1) = f(1) = \frac{1}{6}, F(2) = \frac{2}{6}, \ldots, F(6) = 1$.
  $F(x) = 0$ då $x < 1$ och $F(x) = 1$ då $x \geq 6$.
  Om $\lfloor x \rfloor$ är x avrundat *nedåt*, så kan vi skriva

$$F(x) = \begin{cases} 0 & x < 1 \\ \frac{\lfloor x \rfloor}{6} & 1 \leq x \leq 6 \\ 1 & x > 6. \end{cases}$$

- Ex. $X =$ prickar vid tärningskast.
  $f(1) = f(2) = f(3) = f(4) = f(5) = f(6)$. ($f(x) = 0$ då $x \neq 1, 2, 3, 4, 5, 6$.)
- Ex. $Y =$ myntsinglingar till och med första klaven:

$$P(Y = 1) = P(\text{klave}) = \frac{1}{2},\ P(Y = 2) = P(\text{krona,klave}) \overset{\text{ober.}}{=} \frac{1}{2} \cdot \frac{1}{2}$$

$$P(Y = y) = P(\underbrace{\text{krona,}\ldots\text{,krona}}_{y-1 \text{ gånger}}, \text{klave}) = \frac{1}{2^{y-1}} \frac{1}{2} = \frac{1}{2^y}.$$

Väntevärde

"$\approx$ teoretiskt medelvärde"

- Ex. Tärningskast. Medelvärdet av de olika utfallen är $\frac{1+2+3+4+5+6}{6} = 3.5$. Medelvärdet av många tärningskast kan vi inte förvänta oss vara exakt 3.5, men nära 3.5 i någon mening. (Vi ska senare visa varför—detta är Stora Talens Lag).
- Definition: Väntevärdet av X,

$$\mu = E[X] = \sum_x x f(x).$$

- Ex. Så $X =$ antalet prickar vid ett tärningskast har per definition väntevärde $E[X] = 3.5$.

Väntevärde, forts.

- Om h är någon funktion så gäller $E[h(X)] = \sum_x h(x)f(x)$.
- Ex. Tärningskast.

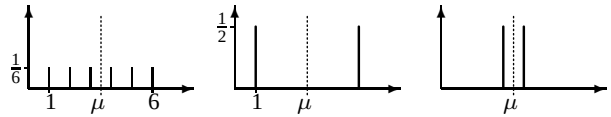
$$h(x) = \left\{ \begin{array}{rl} 100, & x = 6 \\ 0, & x = 5 \\ -20, & x = 1,2,3,4 \end{array} \right\} \Rightarrow$$

$$E[h(X)] = 100 \cdot \tfrac{1}{6} + 0 \cdot \tfrac{1}{6} - 20(\tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6}) = \tfrac{20}{6}.$$

Väntevärdet säger inte allt om en stokastisk variabel

Ex. Tre tärningar

- en vanlig ($X =$ antal prickar vid ett kast),
- en med tre ettor och tre sexor på sina sidor (Y),
- en med tre treor och tre fyror på sina sidor (Z).



Figur: Täthet för X, Y och Z. Alla har samma väntevärde $\mu = 3.5$, men olika spridning kring väntevärdet.

Räkneregler för väntevärdet

a, b, c är konstanter och X och Y två stokastiska variabler.

- $E[c] = c$.
- $E[cX] = cE[X]$.
- $E[X + Y] = E[X] + E[Y]$.
- $E[aX + b] = aE[X] + b$.

Varians

- *Variansen* är ett mått på *spridningen* kring väntevärdet för en stokastisk variabel.
- Definition:

$$\sigma^2 = \mathrm{Var}(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2.$$

- Ex. De tre tärningarna.

$$\begin{aligned} \mathrm{Var}(X) &= E[X^2] - \mu^2 \\ &= \tfrac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) - 3.5^2 \doteq 2.92 \\ \mathrm{Var}(Y) &= \tfrac{1}{6}(1^2 + 6^2) - 3.5^2 = 6.25 \\ \mathrm{Var}(Z) &= \tfrac{1}{6}(3^2 + 4^2) - 3.5^2 = 0.25 \end{aligned}$$

Räkneregler för variansen

a, b, c är konstanter och X och Y två stokastiska variabler.

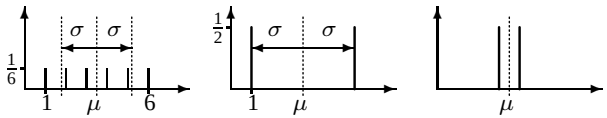
- $\mathrm{Var}(c) = 0$.
- $\mathrm{Var}(cX) = c^2\mathrm{Var}(X)$.
- $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$, *om X och Y är oberoende.*
- $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$.

Bernoulliförsök

- Gör ett ”försök” som lyckas med sannolikhet p. Om det lyckas, låt $I = 1$, om det misslyckas $I = 0$.
- $P(I = 1) = f(1) = p, f(0) = 1 - p$, $E[I] = 1 \cdot f(1) = p$, $\mathrm{Var}(I) = 1^2 \cdot f(1) - p^2 = p - p^2 = p(1 - p)$.
- Detta är den enklaste möjliga stokastiska variabeln, som används som ”byggsten” till flera olika fördelningar.

Standardavvikelse

- Om X mäts i enheten E (poäng, meter, liter/sekund, etc.), så mäts $\mathrm{Var}(X)$ i enheten E^2.
- Därför definieras *standardavvikelsen*: $\sigma = \mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}$, som mäts i enheten E.
- Ex. Tärningarna igen: $\mathrm{SD}(X) \doteq \sqrt{2.92} \doteq 1.71$, $\mathrm{SD}(Y) = \sqrt{6.25} = 2.5$, $\mathrm{SD}(Z) = \sqrt{0.25} = 0.5$.



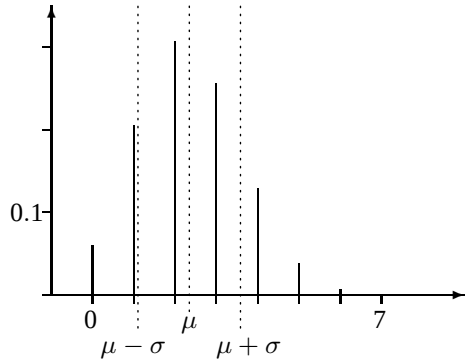
Binomialfördelning och geometrisk fördelning

- Upprepa en följd Bernoulliförsök, som lyckas oberoende av varandra, var och en med sannolikhet p.
- Definition: $X =$ antal lyckade försök bland de n första försöken, sägs vara *binomialfördelad* med parametrar n och p. Vi skriver $X \sim \mathrm{Bin}(n, p)$.
- Definition: $Y =$ antal försök till och med det första lyckade försöket, sägs vara *geometriskt fördelad* med parameter p. $Y \sim \mathrm{Geo}(p)$.

Binomialfördelningen

- Antalet sätt att välja k av n försök som lyckas är $\binom{n}{k}$.
- p^k är sannolikheten för k lyckade försök.
- $(1-p)^{n-k}$ är sannolikheten för $n-k$ misslyckade försök.
- Detta ger $f(k) = \binom{n}{k} p^k (1-p)^{n-k}$.
- $E[X] = np$ (rimligt?) och $\mathrm{Var}(X) = np(1-p)$ visar vi senare med momentgenererande funktioner.
- Det finns tabeller för $F(k)$ för olika kombinationer av n, p, k så att man beräkna t.ex.

$$P(3 \leq X \leq 5) = P(X \leq 5) - P(X \leq 2) = F(5) - F(2).$$



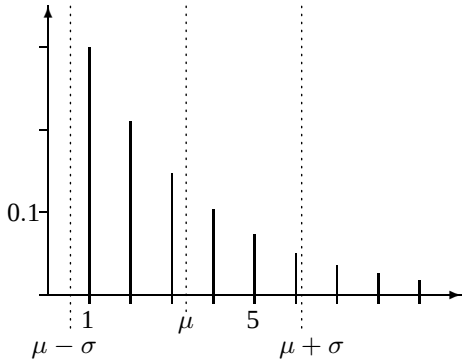
Figur: $f(x)$ för Bin(7,1/3).

Geometrisk fördelning

$$\begin{aligned} f(k) &= P(Y = k) \\ &= P(\text{misslyckas } k-1 \text{ gånger, därefter lyckas}) \\ &= (1-p)^{k-1} p. \end{aligned}$$

$$E[Y] = \frac{1}{p}. \quad \text{(rimligt?)}$$

$$\mathrm{Var}(Y) = \frac{1-p}{p^2}.$$



Figur: $f(x)$ för Geo(0.3).

Exempel

Sannolikheten att Andreas gör minst ett fel under en lektion är 0.3 (han är optimist!). Vad är sannolikheten att han gör

1. fel under precis 5 av 14 lektioner?
2. fel på mellan 3 och 7 av 14 lektioner?
3. första lektionen med ett fel sker på femte lektionen?

Om X är antalet lektioner med fel och Y är antalet lektioner tills första felet dyker upp så är $X \sim \text{Bin}(14, 0.3)$ och $Y \sim \text{Geo}(0.3)$.

1. Svar: $P(X = 5) = \binom{14}{5} 0.3^5 (1 - 0.3)^{14-5} \doteq 0.196$.
2. Svar: $P(3 \leq X \leq 7) = P(X \leq 7) - P(X \leq 2) = 0.9685 - 0.1608 = 0.808$ (enl. tabell)
3. Svar: $P(Y = 5) = (1 - 0.3)^4 \cdot 0.3 = 0.072$.