

Föreläsningar L9MA30, LGMA30

Matematik 3 för lärare, åk 7-9, Sannolikhetslära och statistik,
Matematik 3 för gymnasielärare, Sannolikhetslära och statistik

19 oktober 2017

Föreläsning 1: 1, 4.1-4.2

Sannolikhetslära

Ämnen:

- Vara förtrogen med sannolikhetsbegreppet: beskriva och räkna med (idag)
- Slumpvariabler: sannolikhetsfördelningar

1 Vad är sannolikhetslära?

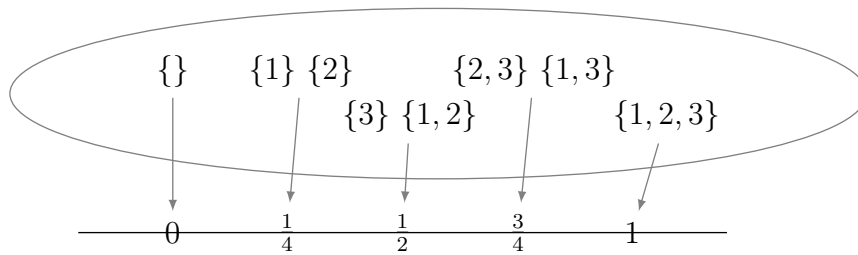
Experiment (eller försök) med en mängd tänkbara resultat,
men vi vet inte vilket av dessa kommer att inträffa.

Ett tänkbart resultat kallas utfall.

$\Omega = \{1, 2, 3\}$ Mängden av alla möjliga utfall kallas utfallsrummet.

$\mathbb{P}(\{1\}) = \frac{1}{4}$ Varje utfall i utfallsrummet är förknippat med ett numeriskt värde
som kallas sannolikhet.

Frågeställning: Vilka saker kan hända? Vad är sannolikheten för dem?



Sannolikhet är ett mått på hur troligt utfallet är!

Samla utfall, händelse och sannolikhet i ett rum: Sannolikhetsrum $(\Omega, \mathcal{F}, \mathbb{P})$, där $\mathcal{F}$ σ -algebra i Ω . Så, får man $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$.

Exempel 1.1. Vi singlar en (symmetrisk) slant: möjliga utfall: krona och klave

$$\Rightarrow \Omega = \{\text{krona, klave}\}$$

med lika stora sannolikheter. Desutom, måste sannolikheterna summera till 1:

$$\Rightarrow \mathbb{P}(\{\text{krona}\}) + \mathbb{P}(\{\text{klave}\}) = 1$$

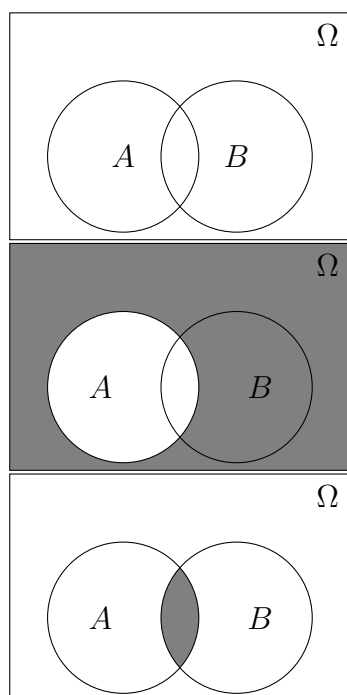
$$\Rightarrow \mathbb{P}(\{\text{krona}\}) = \mathbb{P}(\{\text{klave}\}) = \frac{1}{2}$$

Exempel 1.2. Skolan ska byggas om och det finns två förslag, ett från Emils och ett från Emilias Arkitektbyrå. 100 elever valdes **slumpmässigt** ut. 45 föredrar Emils och 55 Emilias förslag. Vi frågar en slumpvis utvald elev om vilket förslag hen föredrar. Vilket svar kan vi förvänta oss?

Använder den klassiska sannolikhetsdefinitionen:

$$\begin{aligned} \mathbb{P}(\{\text{Emils}\}) &= \frac{\text{antal gynsamma utfall}}{\text{antal möjliga utfall}} \\ &= \frac{45}{100} = 0.45 = p \\ \mathbb{P}(\{\text{Emilias}\}) &= 1 - p = 0.55 \end{aligned}$$

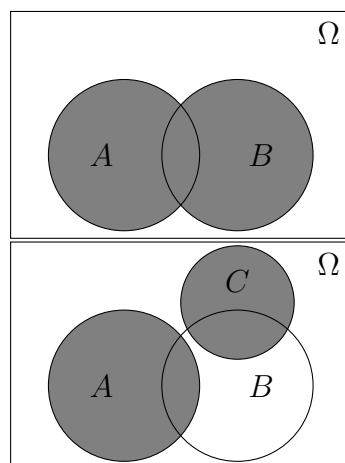
2 Mängdlära



Ett utfallsrum Ω med händelserna A och B

T.ex.: $\Omega = \{\text{kort i en lek}\}$, $A = \{\text{hjärter (13)}\}$, $B = \{\text{ess (4)}\}$, $C = \{\text{ess, ruter kung}\}$

$A^c = A^* = \Omega \setminus A$
t.ex.
 $\{\text{ruter, spader, klöver}\}$



$A \cup B$ "eller"
t.ex.
alla hjärter och
ess

$A \cap B$ "och"
t.ex. hjärter ess

$A \cap C = \emptyset$,
oförenliga,
disjunkta

Definition 2.1. Sannolikhet (Kolmogorov: Axiomsystem)

- $0 \leq \mathbb{P}(A) \leq 1$ för alla händelse A
- $\mathbb{P}(\Omega) = 1$
- Om $A \cap B = \emptyset$ så gäller att $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

Sats 2.1. Sannolikheten $\mathbb{P}$ uppfyller följande påståenden för två godtyckliga händelser A och B i Ω :

- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
- $\mathbb{P}(\emptyset) = 0$
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
- $\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$

Föreläsning 2: 4.3

Rep:

- Nya begrepp: utfall, utfallsrum, händelse, sannolikhet
- Viktiga metoder: Klasiska sannolikhetsdefinitionen: gynsamma/möjliga, venndiagram ex.: 4.5
- Viktiga regler: För godtyckliga händelser A och B
 - $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
 - $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \underbrace{\mathbb{P}(A \cap B)}_{0 \text{ för } A, B \text{ disj.}}$

Idag: Oberoende och betingade sannolikheter, lagen om total sannolikhet, Bayes sats

3 Oberoende händelser

Definition 3.1. Oberoende händelser Två händelser A och B sägs vara oberoende om

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

Exempel 3.1. Vi kastar en tärning och får antingen 1, 2, 3, 4, 5, 6 som disjunkta händelser. Det innebär att om vi får 1, så kan inget annat värde inträffa samtidigt. Det betyder att händelserna är “beroende” av varandra.

$$\begin{aligned}\mathbb{P}(\{1 \text{ på T}\} \cap \{2 \text{ på T}\}) &= \mathbb{P}(\emptyset) = 0 \\ &\neq \mathbb{P}(\{1 \text{ på T}\})\mathbb{P}(\{2 \text{ på T}\}) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}\end{aligned}$$

Om vi nu kastar två tärningar säger vi att de är “oberoende”, för sannolikheten för vilket värde vi får på den ena påverkas inte av utfallet på den andra:

Tärning 1 (T_1) kan ge 1, 2, 3, 4, 5, 6, samma för T_2 :

$$\begin{aligned}\mathbb{P}(\{1 \text{ på } T_1\} \cap \{2 \text{ på } T_2\}) &= \frac{1}{36} \\ &= \frac{1}{6} \cdot \frac{1}{6} = \mathbb{P}(\{1 \text{ på } T_1\})\mathbb{P}(\{2 \text{ på } T_2\})\end{aligned}$$

OBS! Oberoende är inte detsamma som disjunkta!

1	1	2	3	4	5	6
2	1/36	1/36	...			
3	1/36	...				
4	⋮					
5						

Exempel 3.2. Anta $\mathbb{P}(A) = 0.25$, $\mathbb{P}(B) = 0.12$, $\mathbb{P}(A \cup B) = 0.3$. Är A och B oberoende?
Vi vet att

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B) = 0.25 + 0.12 - 0.3 = 0.07$$

men $\mathbb{P}(A)\mathbb{P}(B) = 0.25 \cdot 0.12 = 0.03$ och därmed $\mathbb{P}(A \cap B) \neq \mathbb{P}(A)\mathbb{P}(B) \Rightarrow$ ej oberoende!

4 Betingade sannolikheter

Exempel 4.1. Vi kastar två tärningar igen, men nu betraktar vi summan av prickarna.
Frågeställning hittills:

Vad är sannolikheten för summan 6?

Totalt 36 utfall varav 5 (1+5, 2+4, 3+3, 4+2, 5+1) ger summan 6. $\Rightarrow p = \frac{5}{36}$.

Nu: Givet att den första tärningen ger en 2:a, vad är sannolikheten för summan 6?

Måste få 5:a med $p = \frac{1}{6}$, men

$$\mathbb{P}(\{5 \text{ på } T_2\} | \{2 \text{ på } T_1\}) = ?$$

Intuitivt ansats: $\mathbb{P}(A|B)$

Ex.: Öv. 4.4 40% of patienter med hjärtbesvär är alkoholister:

$B = \{\text{patienter med hjärtbesvär}\}$, $A = \{\text{alkoholister}\} \Rightarrow \mathbb{P}(A|B) = 0.4$

Givet att B inträffar så "händer allt i B ", d.v.s. B är "det nya utfallsrummet"

Hur stor andel av B utgörs av $A \cap B$?

Tillbaka till tärning exemplet:

$$\mathbb{P}(\{5 \text{ på } T_2\} | \{2 \text{ på } T_1\}) = \frac{\mathbb{P}(\{5 \text{ på } T_2\} \cap \{2 \text{ på } T_1\})}{\mathbb{P}(\{2 \text{ på } T_1\})} = \frac{1/36}{1/6} = \frac{1}{6}$$

Definition 4.1. Betingad sannolikhet Låt A och B sägs vara händelser sådana att $\mathbb{P}(B) > 0$. Den betingade sannolikheten för att händelse A skall inträffa givet att B har inträffat skrivs som $\mathbb{P}(A|B)$ och definieras som

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \quad (4.1)$$

OBS! OM A och B är oberoende händelser, så gäller $\mathbb{P}(A|B) = \mathbb{P}(A)$.

Bevis:

$$\mathbb{P}(A|B) \stackrel{\text{def.}}{=} \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \stackrel{\text{obero.}}{=} \frac{\mathbb{P}(A)\mathbb{P}(B)}{\mathbb{P}(B)} = \mathbb{P}(A)$$

Betingningen påverkas inte!

□

5 Lagen om total sannolikhet

Sats 5.1. Lagen om total sannolikhet Låt $A, B_1, B_2, \dots, B_n$ vara händelser sådana att $B_1, \dots, B_n$ delar upp utfallsrummet, d.v.s. är alla disjunkta och utgör tillsammans hela utfallsrummet ($B_i \cap B_j = \emptyset \ \forall i \neq j$ och $B_1 \cup \dots \cup B_n = \Omega$). Då gäller lagen om total sannolikhet:

$$\mathbb{P}(A) = \mathbb{P}(B_1)\mathbb{P}(A|B_1) + \dots + \mathbb{P}(B_n)\mathbb{P}(A|B_n)$$

Bevis för $n = 1, 2$:

Vi vet att (4.1) $\Rightarrow \mathbb{P}(A \cap B) = \mathbb{P}(B)\mathbb{P}(A|B)$

$n = 2$:

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}(A \cap \Omega) = \mathbb{P}(A \cap (B_1 \cup B_2)) \\ &= \mathbb{P}((A \cap B_1) \cup (A \cap B_2)) \\ &\stackrel{B_1 \cap B_2 = \emptyset}{=} \mathbb{P}(A \cap B_1) + \mathbb{P}(A \cap B_2) \\ &\stackrel{'n=1'}{=} \mathbb{P}(B_1)\mathbb{P}(A|B_1) + \mathbb{P}(B_2)\mathbb{P}(A|B_2) \end{aligned}$$

□

Exempel 5.2. Sven spelar på Casino. Han väljer slumpmässigt mellan roulette och blackjack. Han skattar att han vinner i roulette med sannolikhet 0.03 och i blackjack med 0.01. Vad är sannolikhet att han vinner något alls?

Låt $A = \{\text{Sven vinner}\}$, $B_1 = \{\text{Sven spelar roulette}\}$, $B_2 = \{\text{Sven spelar blackjack}\}$.

Sven måste välja ett spel ($B_1 \cup B_2 = \Omega$) men bara ett ($B_1 \cap B_2 = \emptyset$).
Låt $\mathbb{P}(B_1) = \mathbb{P}(B_2) = 0.5$

$$\Rightarrow \mathbb{P}(A) \stackrel{LTS}{=} \mathbb{P}(B_1)\mathbb{P}(A|B_1) + \mathbb{P}(B_2)\mathbb{P}(A|B_2) \\ 0.5 \cdot 0.03 + 0.5 \cdot 0.01 = 0.02$$

Träddiagram:

Exempel 5.3. Sven spelar (igen). Antar att Sven vinner (på något av spelen). Vad är sannolikheten att han spelade roulette?

→ Söker $\mathbb{P}(B_1|A)$ “Vända på betingningen”:

$$\begin{aligned} \mathbb{P}(B_1|A) &\stackrel{def.}{=} \frac{\mathbb{P}(B_1 \cap A)}{\mathbb{P}(A)} \\ &\stackrel{Ex.4}{=} \frac{\mathbb{P}(B_1 \cap A)}{\mathbb{P}(B_1)\mathbb{P}(A|B_1) + \mathbb{P}(B_2)\mathbb{P}(A|B_2)} \\ &\stackrel{bet.slh}{=} \frac{\mathbb{P}(B_1)\mathbb{P}(A|B_1)}{\mathbb{P}(B_1)\mathbb{P}(A|B_1) + \mathbb{P}(B_2)\mathbb{P}(A|B_2)} \\ &= \frac{0.03 \cdot 0.5}{0.02} = 0.75 \end{aligned}$$

Det är tre gånger troligare att vinsten kommer från roulette än från blackjack.

Sats 5.4. Bayes sats Under samma villkor som i lagen om total sannolikhet gäller följande för $i = 1, \dots, n$:

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(B_i)\mathbb{P}(A|B_i)}{\mathbb{P}(B_1)\mathbb{P}(A|B_1) + \dots + \mathbb{P}(B_n)\mathbb{P}(A|B_n)}$$

Bevis:

$$\mathbb{P}(B_i|A) \stackrel{def.}{=} \frac{\mathbb{P}(B_i \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A \cap B_i)}{\mathbb{P}(A)} \stackrel{bet.slh}{=} \frac{\mathbb{P}(B_i)\mathbb{P}(A|B_i)}{\mathbb{P}(A)}$$

□

Föreläsning 3: 4.4

Rep:

- Oberoende händelser: $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$
- Betingad sannolikhet: $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \stackrel{\text{obero.}}{=} \mathbb{P}(A)$
- Lagen om total sannolikhet: $\mathbb{P}(A) = \mathbb{P}(B_1)\mathbb{P}(A|B_1) + \dots + \mathbb{P}(B_n)\mathbb{P}(A|B_n)$
- Bayes sats: $\mathbb{P}(B_i|A) = \frac{\mathbb{P}(B_i)\mathbb{P}(A|B_i)}{\mathbb{P}(B_1)\mathbb{P}(A|B_1) + \dots + \mathbb{P}(B_n)\mathbb{P}(A|B_n)}$
- Ex.: Öv. 4.14 & 15:
 $\mathbb{P}(A \cap B^c) = \mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(A \cap B) = \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A)\mathbb{P}(B^c)$

Idag: Slumpvariabler: Väntevärde, varians

6 Begreppet slumpvariabel; stokastisk variabel

Kom ihåg matematiska funktioner t.ex.:

$$f: \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = x^2, \quad x \in \mathbb{R}$$

NU:

$$X: \Omega \rightarrow \mathbb{R}, \quad \underbrace{X(\omega)}_{\text{fkt. med slumpmässiga värden}} = ?, \quad \omega \in \Omega$$

Exempel 6.1. Låt X vara slant sidan som ligger uppe med $X(\{\text{krona}\})$, $X(\{\text{klave}\})$:
 “Beteckning för alla möjliga utfall av ett experiment/försök/projekt.

Exempel 6.2. Låt X beteckna antalet prickar vi får på en tärning efter vi kastade den:
 $X(\{1\}), \dots, X(\{6\})$

OBS! Innan något försök vet vi inte vad $X(\omega) = X$ är; så X är ingen vanlig variabel, utan en stokastisk variabel. Det enda vi kan beräkna är sannolikheten för olika utfall $\omega \in \Omega$.

Men hur? Vi behöver $\mathbb{P}(X(\omega)) \forall \omega \in \Omega$, d.v.s. funktion eller mått som beskriver slumpstrukturen. Det finns två fall:

- 1) diskreta stokastiska variabler som bara kan anta ett ändligt (eller uppräknligt oändligt (t.ex. alla pos. heltal)) antal värden, t.ex. $X \in \{1, 2, \dots, 6\}$ som för tärningen.

Utfallet i samband med diskreta stokastiska variabler betecknas ofta med k
 $\rightarrow$ IDAG

- 2) kontinuerliga stokastiska variabler som antar ett oändligt (eller **ej** uppräknligt oändligt (t.ex. alla pos. heltal)) antal värden, t.ex. $X \in [0, \infty)$ väntetid på Lotteri vinst.

Utfallet i samband med kontinuerliga stokastiska variabler betecknas ofta med små bokstäver, t.ex. x
 $\rightarrow$ NÄSTA GÅNG

Definition 6.1. Sannolikhetsfunktion för diskreta X

$$p_X(k) = \mathbb{P}(X = k) = \mathbb{P}(X \text{ antar värdet } k)$$

som uppfyller $p_X(k) \geq 0$ och $\sum_k p_X(k) = 1$.

Exempel 6.3. Låt X beteckna antalet prickar vi får på en tärning: $p_X(k) = \frac{1}{6}, k = 1, \dots, 6$ och 0 annars.

Definition 6.2. Fördelningsfunktion för diskreta X

$$F_X(x) = \underbrace{\mathbb{P}(X \leq x)}_{\text{kumulativ sannolikhet}} = \sum_{k \leq x} p_X(k), \quad \forall x \in \mathbb{R}$$

Exempel 6.4. Låt X beteckna antalet prickar vi får på en tärning:

$$F_X(x) = \underbrace{\mathbb{P}(X \leq x)}_{\text{kumulativ sannolikhet}} = \begin{cases} 0 & , -\infty < x < 0 \\ \text{heltalsdelen}(x/6) & , 0 \leq x \leq 6 \\ 1 & , 6 < x < \infty \end{cases} = \underbrace{\sum_{k=1}^x \frac{1}{6}}_{\text{för visst } x}$$

Nu och på torsdag: Hur kan man beskriva fördelningen av X på en enklare sätt?

- Väntevärde: Värdet vi kan förvänta oss att få: teoretiskt medelvärde, populationsmedelvärde
- Varians: Hur sprider värden kring väntevärdet?

7 Väntevärde för diskreta s.v.

Definition 7.1. Väntevärde för diskreta X

$$\mathbb{E}(X) = \sum_k k \cdot p_X(k)$$

Exempel 7.1. Låt X beteckna antalet prickar vi får på en tärning:

$$\mathbb{E}(X) = 1 \cdot \frac{1}{6} + \dots + 6 \cdot \frac{1}{6} = 3.5$$

Exempel 7.2. En lott kostar 10kr och man kan vinna 50kr med slh. 0.1 och 100kr med slh. 0.01. Hur mycket kommer lotteriet (i snitt) att tjäna på varje såld lott?

Låt X vara vinsten

$$\mathbb{P}(X = 0) = 1 - 0.1 - 0.01 = 0.89, \quad \mathbb{P}(X = 50) = 0.1, \quad \mathbb{P}(X = 100) = 0.01$$

$$\mathbb{E}(X) = 0 \cdot 0.89 + 50 \cdot 0.1 + 100 \cdot 0.01 = 6[kr]$$

Eftersom lotten kostar 10kr, tjänar lotteriet 4kr per lott i snitt.

Sats 7.3. Räknerregler för väntevärden Låt X och Y vara godtyckliga s.v. och $a, b, c \in \mathbb{R}$ godtyckliga konstanter. Då gäller

$$\mathbb{E}(aX + bY + c) = a\mathbb{E}(X) + b\mathbb{E}(Y) + c$$

och för en funktion g

$$\mathbb{E}(g(X)) = \sum_k g(k)p_X(k)$$

Exempel 7.4. Låt X beteckna antalet prickar vi får på en tärning:

$$\mathbb{E}(X^2) = 1^2 \cdot \frac{1}{6} + \dots + 6^2 \cdot \frac{1}{6} = 15.17$$

8 Varians för diskreta s.v.

Definition 8.1. Varians för diskreta X Variansen mäter den förväntade quadratiska avvikelsen från väntevärdet:

$$V(X) = \mathbb{E} \left((X - \mathbb{E}(X))^2 \right) = \sum_k (k - \mathbb{E}(X))^2 \cdot p_X(k)$$

OBS! Det finns ett enklare sätt att beräkna variansen:

$$V(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

Exempel 8.1. Låt X beteckna antalet prickar vi får på en tärning:

$$V(X) = 15.17 - (3.5)^2 = 2.92$$

Föreläsning 4: 4.4

Idag (som i mådags): Slumpvariabler: Väntevärde, varians

9 Begreppet slumpvariabel; stokastisk variabel

Innan något försök vet vi inte vad X är så X är ingen vanlig variabel, utan en stokastisk variabel. Det enda vi kan beräkna är sannolikheten för olika händelser $\subset \Omega$.

Men hur? Vi behöver $\mathbb{P}(A) \forall A \subset \Omega$, d.v.s. funktion eller mått som beskriver slumpstrukturen. Det finns två fall:

- 1) diskreta stokastiska variabler som bara kan anta ett ändligt (eller uppräknligt oändligt (t.ex. alla pos. heltal)) antal värden, t.ex. $X \in \{1, 2, \dots, 6\}$ som för tärningen.

Utfallet i samband med diskreta stokastiska variabler betecknas ofta med k
→ i mådags: T.ex. X = antalet prickar på en tärning

- 2) kontinuerliga stokastiska variabler som antar ett oändligt (eller **ej** uppräknligt oändligt (t.ex. alla pos. heltal)) antal värden, t.ex. $X \in [0, \infty)$ väntetid på Lotteri vinst.

Utfallet i samband med kontinuerliga stokastiska variabler betecknas ofta med små bokstäver, t.ex. x
→ IDAG: T.ex. X = väntetid på bussen som går var femte minut (alldrig för sent) när man går till hållplatsen "på vinst och förlust"

Definition 9.1. Rep: Sannolikhetsfunktion för diskreta X

$$p_X(k) = \mathbb{P}(X = k) = \mathbb{P}(X \text{ antar värdet } k)$$

som uppfyller $p_X(k) \geq 0$ och $\sum_k p_X(k) = 1$.

Definition 9.2. Ny: Täthetsfunktion för kontinuerlig X En funktion f_X som uppfyller $f_X(x) \geq 0$ och $\int_{\mathbb{R}} f_X(x) dx = 1$ för alla $x \in \mathbb{R}$ kallas en täthetsfunktion.

OBS! $\mathbb{P}(X = x) = \frac{1}{\infty} = 0 \neq f_X(x)$.

Definition 9.3. Fördelningsfunktion för s.v.

$$\begin{aligned}
 F_X(x) &= \underbrace{\mathbb{P}(X(k) \leq x)}_{\text{kumulativ sannolikhet}}, \quad \forall x \in \mathbb{R} \\
 &\stackrel{\text{disk.}}{=} \sum_{k \leq x} p_X(k) \\
 &\stackrel{\text{kont.}}{=} \int_{-\infty}^x f_X(t) dt
 \end{aligned}$$

Exempel 9.1. Ny: Låt X beteckna väntetid på bussen: Vi vet att $\int_0^5 f_X(x) dx = 1$ och därmed

$$f_X(x) = \begin{cases} \frac{1}{5} & , 0 \leq x \leq 5 \\ 0 & , \text{annars} \end{cases}$$

och

$$F_X(x) = \begin{cases} 0 & , x < 0 \\ \frac{x}{5} & , 0 \leq x \leq 5 \\ 1 & , 5 > x \end{cases}$$

10 Väntevärde för s.v.**Definition 10.1. Väntevärde för s.v.**

$$\begin{aligned}
 \mathbb{E}(X) &\stackrel{\text{disk.}}{=} \sum_k k \cdot p_X(k) \\
 &\stackrel{\text{kont.}}{=} \int_{\mathbb{R}} x \cdot f_X(x) dx
 \end{aligned}$$

Vi kan tänka på väntevärdet som medelvärde av n observationer då $n \rightarrow \infty$. Alternativt: tyngdpunkten av fördelningen. Om $f_X(x)$ hade betecknat densitet av en kropp, hade $\mathbb{E}[X]$ varit kroppens tyngdpunkt (i x -led).

Sats 10.1. Räknerregler för väntevärden Låt X och Y vara godtyckliga s.v. och $a, b, c \in \mathbb{R}$ godtyckliga konstanter. Då gäller

$$\mathbb{E}(aX + bY + c) = a\mathbb{E}(X) + b\mathbb{E}(Y) + c$$

och för en funktion g

$$\begin{aligned} \mathbb{E}(g(X)) &\stackrel{\text{disk.}}{=} \sum_k g(k)p_X(k) \\ &\stackrel{\text{kont.}}{=} \int_{\mathbb{R}} g(x)f_X(x)dx \end{aligned}$$

Sats 10.2. Väntevärdet för ett medelvärde Låt $X_1, \dots, X_n$ ha samma väntevärde μ ($\mathbb{E}(X_i) = \mu \quad \forall i$) och beteckna medelvärdet med $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$. Då gäller

$$\mathbb{E}(X) = \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \stackrel{\text{reg.}}{=} \frac{1}{n} \sum_{i=1}^n \underbrace{\mathbb{E}(X_i)}_{\mu} = \frac{1}{n} \cdot n \cdot \mu = \mu$$

11 Varians för s.v.

För att ytterligare karaktärisera s.v.: T.ex. dessa här har samma väntevärde μ . Hur kan vi beskriva skillnaden?

Exempel 11.1. Antag att vi har två bussar att välja på. Den ena går härifrån var tionde minut, men vi vet inte när. Låt tiden tills vi kan sitta på den bussen vara X . Den andra går var femte minut, men vi vet inte när och det tar 2.5 minuter att gå till hållplatsen. Låt tiden tills vi kan sitta på den andra på bussen vara Y . Notera:

$$\begin{aligned} f_X(x) &= 1/10 \quad 0 \leq x \leq 10 \\ f_Y(y) &= 1/5 \quad 2.5 \leq y \leq 7.5 \end{aligned}$$

$\mathbb{E}[X] = \mathbb{E}[Y] = 5$, vad är skillnaden?

Theorem and Definition 11.1. Varians för s.v. Variansen mäter den förväntade kvadratiske avvikelsen från väntevärdet:

$$V(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

Bevis:

$$\begin{aligned}
 V(X) &\stackrel{\text{def.}}{=} \mathbb{E} \left((X - \mathbb{E}(X))^2 \right) \\
 &= \mathbb{E} \left(X^2 - 2X\mathbb{E}(X) + (\mathbb{E}(X))^2 \right) \\
 &\stackrel{\text{reg.}}{=} \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \underbrace{\mathbb{E}((\mathbb{E}(X))^2)}_{\text{konst.}} \\
 &\stackrel{\text{reg.}}{=} \mathbb{E}(X^2) - 2(\mathbb{E}(X))^2 + (\mathbb{E}(X))^2 \\
 &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2
 \end{aligned}$$

□

Exempel 11.2. Låt X beteckna väntetid på första bussen:

$$\begin{aligned}
 \mathbb{E}(X) &= \int_0^{10} x \cdot \frac{1}{10} dx = \left[\frac{1}{10} \frac{x^2}{2} \right]_0^{10} = 5 \\
 \mathbb{E}(X^2) &= \int_0^{10} x^2 \cdot \frac{1}{10} dx = \left[\frac{1}{10} \frac{x^3}{3} \right]_0^{10} = \frac{100}{3} \approx 33.333 \\
 V(X) &= 33.333 - (5)^2 = 8.3333
 \end{aligned}$$

Sats 11.3. Räkner regler för varianser Låt X och Y vara oberoende s.v. och $a, b, c \in \mathbb{R}$ godtyckliga konstanter. Då gäller

$$V(aX + bY + c) = a^2V(X) + b^2V(Y)$$

Bevis: Visa att

- 1) $V(aX + b) = a^2V(X) \longrightarrow$ *dugga*
- 2) $V(X+Y) = V(X) + V(Y)$

Exempel 11.4. Låt Y beteckna tid till andra bussen, och låt $Z = Y - 2.5$ (alltså väntetiden när vi kommit till hållplatsen). Då gäller: $V(Y) = V(Z + 2.5) = V(Z)$.

$$\begin{aligned}
 \mathbb{E}(Z) &= 2.5 \\
 \mathbb{E}(Z^2) &= \int_0^5 z^2 \cdot \frac{1}{5} dz = \left[\frac{1}{5} \frac{z^3}{3} \right]_0^5 = \frac{25}{3} \approx 8.333 \\
 V(Z) &= 8.3333 - (2.5)^2 = 2.083
 \end{aligned}$$

Definition 11.1. Standardavvikelsen

$$D(X) = \sqrt{V(X)} \quad (\text{standardisering})$$

OBS! Om s.v. X har enheten m , så har $V(X)$ enheten m^2 och $D(X)$ enheten m och därmed “fysikalisk tolkning”. T.ex.: Lott

$$V(X) = 314kr^2 \quad \text{och} \quad D(X) = \sqrt{314kr^2} = 17.72kr$$

Sats 11.5. Varians och standardavvikelse för ett medelvärde Låt $X_1, \dots, X_n$ vara oberoende och ha samma standardavvikelse σ ($D(X_i) = \sigma \quad \forall i$) och beteckna medelvärdet med $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$. Då gäller

$$\begin{aligned} \text{Var}(X) &= \text{Var} \left(\frac{1}{n} \sum_{i=1}^n X_i \right) \stackrel{\text{reg.}}{=} \frac{1}{n^2} \sum_{i=1}^n \underbrace{\text{Var}(X_i)}_{\sigma} = \frac{n}{n^2} \cdot \sigma^2 = \frac{1}{n} \cdot \sigma^2 \\ D(X) &= \frac{\sigma}{\sqrt{n}} \end{aligned}$$

Föreläsning 5: 4.5

Rep.: Kontinuerliga s.v.

- Fördelningsfkt.: $F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x \underbrace{f_X(t)}_{\text{täthetsfkt.}} dt$
- $\mathbb{E}(X) = \int_{\mathbb{R}} x f_X(x) dx$
- $V(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$
- $D(X) = \sqrt{V(X)}$

12 Några diskreta sannolikhetsfördelningar

Motiv: Hittills har vi mest antagit att alla utfall är lika sannolika, som i tärningsexemplen. Detta är exempel *likformiga* fördelningar.

Definition 12.1. Likformig (diskret) fördelning En diskret s.v. X är likformigt fördelad, vilket skrivs som $X \sim U(N)$ om

$$p_X(k) = \begin{cases} \frac{1}{N} & , 1 \leq k \leq N \\ 0 & , \text{annars} \end{cases}$$

Vårt favoritexempel, antalet prickar en tärning visar, är en likformigt fördelad slumpvariabel. Det skulle också slantsinglingarna kunna vara, men det finns även ett annat sätt att betrakta dem på. Antag att vi istället kallar de olika utfallen 0 och 1, alltså till exempel 0 = "krona" och 1 = "klave". Antag vidare att vi har ett (möjligen) asymmetriskt mynt, så att sannolikheten att få krona är $p \in [0, 1]$. Då har vi följande fördelning:

Definition 12.2. Bernoulli-fördelning En s.v. X är Bernoulli-fördelad, vilket skrivs som $X \sim \text{Bern}(p)$ om

$$p_X(k) = \begin{cases} p & , k = 1 \\ 1 - p & , k = 0 \\ 0 & , \text{annars} \end{cases}$$

Idag och på torsdag ska vi se fyra andra fördelningar: Binomial, hypergeometrisk, Poisson, för första gången.

Repetition: $\mathbb{E}[X] = \sum_k k p_X(k)$ Till exempel:

$$X \sim U(N) \Rightarrow \mathbb{E}[X] = \sum_{k=1}^N k/N = \frac{1}{N} \frac{N(N+1)}{2} = \frac{N+1}{2}$$

$$X \sim \text{Ber}(p) \Rightarrow \mathbb{E}[X] = 1 \cdot p + 0 \cdot (1 - p) = p$$

$V(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$. Till exempel:

$$\begin{aligned} X \sim U(N) &\Rightarrow V(X) = \frac{(N+1)(N-1)}{12} \\ X \sim \text{Ber}(p) &\Rightarrow V(X) = p(1-p) \end{aligned}$$

Exempel 12.1. Uppgift 18 En fågel föder ungar. Antalet och sannolikheterna ges enligt tabell:

k	k^2	$p_X(k)$	$F_X(k)$
0	0	0.3	0.3
1	1	0.35	0.65
2	4	0.25	0.9
3	9	0.1	1

Beräkna $p_X(k)$, $F_X(x)$, $\mathbb{E}[X]$, $V(X)$ samt $D(X)$. (Tabell, Ny tabell, 1.15 , $2.25 - 1.15^2 = 0.9275$, 0.9630).

Definition 12.3. Binomialkoefficient Antalet *kombinationer* (dvs utan hänsyn till ordning) av k element ur n skrivs

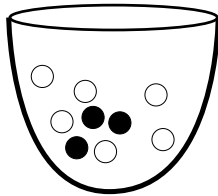
$${}_nC_k = \binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{1}{k!} \underbrace{{}_nP_k}_{\text{permutationer}}$$

Konvention: $0! = 1$

Hur förstår vi detta!? Vi börjar med att kolla antalet sätt vi kan ordna saker. Antalet sätt vi kan ordna n element på är $n!$. Men nu ska vi bara välja k ur n , så det första kan vi välja på n sätt, det andra på $(n-1)$... på plats nummer k har $(n-k+1)$ kvar att välja på. Antalet *permutationer* är alltså $n(n-1)(n-2)\dots(n-k+1)$. Men detta är precis $\frac{n!}{(n-k)!}$. Nu ville vi ha antalet *kombinationer*, så ordningen på de k element vi väljer ska inte spela någon roll. Men antalet sätt att ordna k objekt på är $k!$, så antalet permutationer av k ur n måste vara $k! \cdot (\text{antalet kombinationer})!$

12.1 Hypergeometrisk fördelningen

Antag att vi har en urna med totalt N kulor i, varav $m = Np$ är svarta (alltså $p = m/N$). Vi drar nu n kulor ur urnan och låter X = antalet svarta kulor vi drar. Vad är $p_X(k)$? Låt oss anta att vi drar $n = 5$ kulor ur följande urna med $N = 9$, $m = 3$ dvs $p = 1/3$.



Notera att alla kulor har samma sannolikhet att bli dragna, så vi kan använda klassiska sannolikhetsdefinitionen, dvs

$$p_X(k) = \frac{\text{gynnsamma}}{\text{möjliga}} = \frac{\text{antal kombinationer med } k \text{ svarta}}{\text{antal möjliga}}$$

Antal möjliga? $\binom{N}{n} = \binom{9}{5} = 126$. Gynnsamma? Tänk att vi hade haft vita och svarta i separata urnor. Vi hade då kunnat välja

$$\begin{aligned} \text{de svarta på } \binom{Np}{k} &= \binom{3}{k} \text{ sätt, och} \\ \text{de vita på } \binom{N(1-p)}{n-k} &= \binom{6}{n-k} \text{ sätt.} \end{aligned}$$

Restriktioner? Kan inte dra fler svarta än m (eller n såklart). Även: Kan inte dra fler *vita* än $N(1-p)$, och antalet vita vi drar är $n-k$.

Definition 12.4. Hypergeometrisk fördelning En diskret s.v. X är hypergeometriskt fördelad med parametrar N, n och p om

$$p_X(k) = \begin{cases} \frac{\binom{Np}{k} \binom{N(1-p)}{n-k}}{\binom{N}{n}} & 0 \leq k \leq Np \text{ och } 0 \leq n-k \leq N(1-p) \\ 0 & \text{, annars} \end{cases}$$

där $m = Np$ måste vara ett heltal. För X gäller

$$\mathbb{E}[X] = np \quad \text{och} \quad V(X) = np(1-p) \frac{N-n}{N-1}$$

Vad är t ex att vi drar en svart kula i urn-exemplet ovan? $N=9, n=5, p=1/3$ så

$$p_X(1) = \frac{\binom{3}{1} \binom{6}{4}}{\binom{9}{5}} \approx 0.357$$

Vi kan också se $\mathbb{E}[X] = 5\frac{1}{3} = 1.67$ och $V(X) = 5\frac{1}{3}\frac{2}{3}\frac{9-5}{9-1} \approx 0.556 \Rightarrow D(X) \approx 0.745$

Nu: vad händer om det finns oändligt många kulor i urnan? Dvs vi låter $N \rightarrow \infty$ med $p = m/N$ fixt".

12.2 Binomialfördelningen

Antag att vi gör n oberoende Bernoulli-försök med samma p , dvs varje försök lyckas med sannolikhet p och misslyckas med sannolikhet $1 - p$. Ex: singla slant, skjuta på mål, få barn...

Låt $X \in \{0, 1, 2, \dots, n\} = \Omega$ vara antalet lyckade försök. Vad är $p_X(k)$? Låt oss betrakta fallet att alla försök lyckas, alltså $\mathbb{P}(X = n)$. Detta måste vara $p \cdot p \dots p = p^n$. På samma sätt: $\mathbb{P}(X = 0) = (1 - p)^n$. Låt oss nu betrakta en *viss* sekvens med k lyckade försök och $n - k$ misslyckade (t ex de första k lyckas, resten misslyckas). Denna måste ha sannolikhet $p^k(1 - p)^{(n-k)}$. Men det finns flera sådana sekvenser! Hur många...? Svar: $\binom{n}{k}$.

Definition 12.5. Binomialfördelad s.v. En s.v. X är binomialfördelad med parametrar $p \in [0, 1]$ och $n \in \mathbb{N}$, vilket skrivs $X \sim \text{Bin}(n, p)$ om

$$p_X(k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k = 0, 1, \dots, n$$

Då gäller

$$\begin{aligned} \mathbb{E}[X] &= np \text{ och} \\ V(X) &= np(1 - p) \end{aligned}$$

För att härleda väntevärde och varians kan vi använda följande formulering: Låt $Y_1, Y_2, \dots, Y_n \sim \text{Bern}(p)$ och $X = \sum_{i=1}^n Y_i$. Då är $X \sim \text{Bin}(n, p)$.

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}\left[\sum_{i=1}^n Y_i\right] \stackrel{\text{reg.}}{=} \sum_{i=1}^n \mathbb{E}[Y_i] = np \\ V(X) &= V\left[\sum_{i=1}^n Y_i\right] \stackrel{\text{reg.}}{=} \sum_{i=1}^n V[Y_i] = np(1 - p) \end{aligned}$$

Föreläsning 6: 4.5

Rep.: Diskreta fördelningar

1 Likformig, $X \sim U(N)$:

$$\begin{aligned} p_X(k) &= 1/N, \quad k = 1, 2, \dots, N \\ \mathbb{E}[X] &= \frac{N+1}{2} \\ V(X) &= \frac{(N+1)(N-1)}{12} \end{aligned}$$

2 Bernoulli: $X \sim Ber(p)$

$$\begin{aligned} p_X(1) &= p; \quad p_X(0) = 1 - p \\ \mathbb{E}[X] &= p \\ V(X) &= p(1 - p) \end{aligned}$$

3 Hypergeometrisk fördelningen, $X \sim Hyp(N, n, p)$

$$\begin{aligned} p_X(k) &= \frac{\binom{Np}{k} \binom{N(1-p)}{n-k}}{\binom{N}{n}}, \quad 0 \leq k \leq N(1-p) \text{ och } 0 \leq n-k \leq N(1-p) \\ \mathbb{E}[X] &= np \\ V(X) &= np(1-p) \frac{N-n}{N-1} \end{aligned}$$

4 Binomialfördelningen, $X \sim Bin(n, p)$

$$\begin{aligned} p_X(k) &= \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n \\ \mathbb{E}[X] &= np \\ V(X) &= np(1-p) \end{aligned}$$

Idag: två till, "för första gången" och Poisson.

12.3 För första gången-fördelningen (geometrisk fördelningen)

Antag att vi gör Bernoulli försök, dvs varje försök lyckas eller misslyckas. Men istället för att räkna antalet lyckade (som i binomial-fördelningen), så fortsätter vi tills ett försök lyckas, och låter X = "antal försök tom första lyckade". Då gäller $X \sim ffg(p)$.

Definition 12.6. För första gången-fördelningen. Låt X vara en diskret slumpvariabel. Då är X för första gången-fördelad med parameter p , $X \sim ffg(p)$, om

$$p_X(k) = (1-p)^{k-1}p, \quad k = 1, 2, \dots$$

Då gäller:

$$\mathbb{E}[X] = 1/p, \quad \text{och} \quad V(X) = (1-p)/p^2$$

Notera: i boken säger dom att geometriska fördelningen är lite annorlunda (räknar antalet *misslyckanden* innan lyckas). Jag håller inte med, tycker geometriska och ffg-fördelningarna är samma.

12.4 Poisson-fördelningen

Poisson-fördelningen beskriver diskreta händelser i kontinuerlig tid (eller area, volym eller liknande). Antag t ex att vi betraktar kunder som kommer till en butik under en timme, och att det i genomsnitt kommer λ kunder /h. En grov uppskattning skulle ges av att vi för varje minut bara registrerar om det kommit någon kund eller ej, och summerar dessa. Då gör vi $n = 60$ oberoende försök, med sannolikhet p innehåller en minut en kund. Så $X = \text{“antal kunder”}$ bör vara ungefärligt $\sim \text{Bin}(n, p)$, med $np = \lambda$. Men det kan definitivt komma fler än 1 kund under en minut, så vi borde titta på finare tidsuppdelning! Låt oss ta sekunder istället. Så $n = 3600$, och p nu betydligt lägre för vi behöver fortfarande $np = \lambda$. Det kan såklart komma fler än 1 kund på en sekund också... vad händer om vi låter $n \rightarrow \infty$? Alltså, låt $X \sim \text{Bin}(n, p)$ och låt $n \rightarrow \infty$, samtidigt som vi håller $np = \lambda$ konstant. Vad är $p_X(k)$?

$$\begin{aligned} p_X(k) &= \binom{n}{k} p^k (1-p)^{n-k} = \{p = \lambda/n\} = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{n!}{k!(n-k)!} \frac{\lambda^k (1 - \lambda/n)^n}{n^k (1 - \lambda/n)^k} = \\ &= \frac{n!}{n^k k!(n-k)!} \frac{\lambda^k (1 - \lambda/n)^n}{(1 - \lambda/n)^k} \rightarrow \frac{\lambda^k}{k!} e^{-\lambda} \\ &\quad \left\{ \frac{n!}{n^k k!(n-k)!} \approx n^k/n^k \rightarrow 1, (1 - \lambda/n)^k \rightarrow 1, (1 - \lambda/n)^n \rightarrow \exp(-\lambda) \right\} \end{aligned}$$

Detta är Poisson-fördelningen!

Definition 12.7. Poisson-fördelningen En diskret s.v., X , är Poisson-fördelad med parameter λ om

$$p_X(k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, \dots$$

vilket skriv som $X \sim Poi(\lambda)$. Då gäller

$$\mathbb{E}[X] = \lambda, \text{ och } V(X) = \lambda$$

Exempel 12.2. Ett radioaktivt preparat sönderfaller med i genomsnitt 10 sönderfall per sekund. Vad är sannolikheten att det sker som mest 5 sönderfall under en sekund? (Poisson, ty i diskret tid, inget beroende). Låt X = "antalet sönderfall". Då gäller $X \sim Poi(10)$. Vi söker

$$\mathbb{P}(X \leq 5) = F_X(5) = \text{Tabell sid 390} = 0.0671$$

Exempel 12.3. Kalle skjuter 20 basketstraffar. Han sätter var och en med sannolikhet $p = 0.3$. Vad är sannolikheten att han sätter minst hälften?

Låt X = "antalet straffar Kalle sätter" $\Rightarrow X \sim Bin(20, 0.3)$. Söks $\mathbb{P}(X \geq 10) = 1 - \mathbb{P}(X < 10) = 1 - \mathbb{P}(X \leq 9) = 1 - \underbrace{F_X(9)}_{\text{fördelningsfkt.}}$.

$$F_X(9) = \sum_{k=0}^9 p_X(k) = \sum_{k=0}^9 \binom{20}{k} 0.3^k 0.7^{(20-k)}$$

INTE sugen på att räkna summan. Men tabell på sid. 381-386 säger: $F_X(9) = 0.952$, dvs $\mathbb{P}(X \geq 10) = 1 - 0.952 = 0.048$. Notera även

$$\mathbb{E}[X] = np = 6$$

$$V(X) = np(1-p) = 4.2$$

Exempel 12.4. I en klass finns 32 elever varav 4 är duktiga simmare. En grupp av 8 elever väljs slumpmässigt för en simtävling. Vad är sannolikheten att det finns 2 duktiga simmare i gruppen?

Låt X = "antal duktiga simmare i gruppen", då gäller $X \sim Hyp(32, 4, 1/8)$. Söker $\mathbb{P}(X = 2)$.

$$\mathbb{P}(X = 2) = p_X(2) = \frac{\binom{Np}{2} \binom{N(1-p)}{n-2}}{\binom{N}{n}} = \frac{\binom{4}{2} \binom{32-4}{8-2}}{\binom{32}{8}} = \frac{6 \cdot 376740}{10518300} = 0.215$$

Hur många duktiga simmare kan man förvänta sig i gruppen? $\mathbb{E}[X] = np = 8 \cdot \frac{1}{8} = 1$.

Vi kanske vill veta nåt mer?

$$\mathbb{P}(X \leq 2) = F_X(2) = \sum_{k \leq 2} p_X(k) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) =$$

$$0.295 + 0.450 + 0.215 = 0.96$$

$$\mathbb{P}(X \geq 2) = \sum_{k \geq 2} p_X(k) = \mathbb{P}(X = 2) + \mathbb{P}(X = 3) + \mathbb{P}(X = 4)$$

$$1 - \mathbb{P}(X < 2) = 1 - \mathbb{P}(X \leq 1) = 1 - F_X(1) = 1 - (0.295 + .45) = 0.255$$

$$\mathbb{P}(1 \leq X < 3) = \mathbb{P}(X = 1) + \mathbb{P}(X = 2) = \mathbb{P}(X \leq 2) - \mathbb{P}(X \leq 0) = \\ F_X(2) - F_X(0) = 0.96 - 0.295 = 0.665$$

Föreläsning 7: 4.6

Rep.: Diskreta s.v.

- 1) **Likformig**, $X \sim U(N)$: Alla utfall lika sannolika (tärningskast, kort ur lek, etc).
- 2) **Bernoulli**, $X \sim Ber(p)$: Ett försök, lyckas med sannolikhet p , misslyckas med sannolikehet $1 - p$ (asymmetrisk slant, ett basketskott, vad som helst som lyckas/misslyckas).
- 3) **Hypergeometrisk**, $X \sim Hyp(N, n, p)$: Drar n objekt ur N varav en andel $p = m/N$ har en viss egenskap utan återläggning. Räknar antalet "svarta kulor" (Kulor ur urnor, lotto, antal duktiga simmare, etc).
- 4) **Binomial**, $X \sim Bin(n, p)$: Gör n oberoende försök som lyckas med sannolikhet p och misslyckas med sannolikhet $1 - p$. Räknar antal lyckade försök (flera slantsinglingar, basketskott, etc.)
- 5) **Poisson**, $X \sim Poi(\lambda)$: Diskreta händelser i tid, med givet genomsnitt, λ , per tidsenhet. Räknar antal händelser (kärnsönderfall, kunder till butik, etc).
- 6) **För första gången**, $X \sim ffg(p)$: Gör Bernoulli-försök, räknar antal försök tom första lyckade (singlingar till första "krona", antal försök till första basketskottet sitter, etc).

Idag:

13 Några kontinuerliga fördelningar

13.1 Likformig fördelning

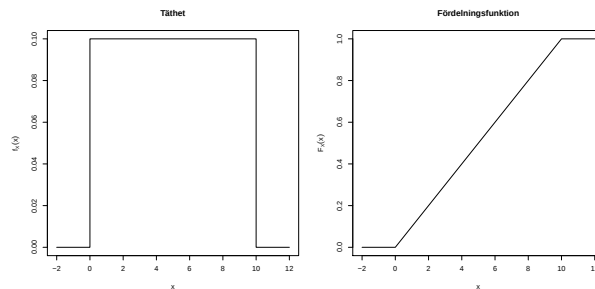
Kom ihåg bussexemplet, vi går på vinst och förlust till hållplatsen för att ta en buss som går var tionde minut och låter $X = \text{"väntetid till bussen"}$.

Vi hade då att

$$f_X(x) = \begin{cases} 1/10 & , \quad 0 \leq x \leq 10 \\ 0 & , \quad \text{annars} \end{cases}$$

samt

$$F_X(x) = \begin{cases} 0 & , \quad x < 0 \\ x/10 & , \quad 0 \leq x \leq 10 \\ 1 & , \quad x > 10 \end{cases}$$



Hur stor är sannolikheten att vi får vänta mer än 7 min?

$$\mathbb{P}(X > 7) = \mathbb{P}(7 \leq X \leq 10) = \mathbb{P}(7 < X < 10) = \int_7^{10} f_X(x) dx = F_X(10) - F_X(7) = 0.3$$

Tolkning av tätheten:

$$f_X(x)\Delta x \approx F_X(x + \Delta x) - F_X(x) = \mathbb{P}(x \leq X \leq x + \Delta x)$$

Nu: Generaliserad.

Definition 13.1. Likformig (kontinuerlig) fördelning En kontinuerlig s.v., X , är likformigt fördelad på intervallet $[a, b]$, vilket skrivs som $X \sim U([a, b])$, om den har täthet

$$f_X(x) = \begin{cases} 1/(b-a) & , \quad a \leq x \leq b \\ 0 & , \quad \text{annars} \end{cases}$$

Då gäller

$$\mathbb{E}[X] = \int_a^b x f_X(x) dx = \frac{1}{b-a} \int_a^b x dx = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}$$

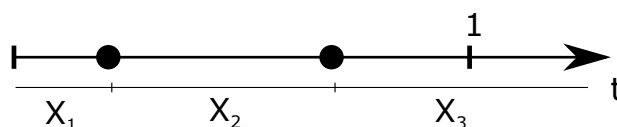
$$V(X) = \frac{(b-a)^2}{12}, \quad \text{LÄXA}$$

13.2 Exponential-fördelningen

Används för att beskrivs livslängder (hos maskiner etc), kötider, osv.

Sammanfattat: tid det tar tills något, vars sannolikhet att inträffa i ett visst intervall bara beror på längden på intervallet, sker.

Exempel 13.1. Låt $Y \sim Poi(\lambda)$ vara antalet sönderfall under en sekund, dvs $p_Y(k) = \frac{\lambda^k}{k!} e^{-\lambda}$. Låt vidare $X_1, X_2, \dots$ vara väntetiderna mellan sönderfall.



Vi kan fundera på sannolikheten att $X_1 \leq 1$.

$$\mathbb{P}(X_1 \leq 1) = 1 - \mathbb{P}(X_1 > 1) = 1 - \mathbb{P}(Y = 0) = 1 - e^{-\lambda}$$

Med samma metod kan man visa att

$$\mathbb{P}(X \leq x) = F_X(x) = 1 - e^{-\lambda x} \Rightarrow f_X(x) = \frac{d}{dx} F_X(x) = \lambda e^{-\lambda x}, \quad \forall x \geq 0$$

Detta är exponentialfördelningen!

Definition 13.2. Exponentialfördelningen En kontinuerlig s.v., X , är exponentialfördelad med parameter β om

$$f_X(x) = \begin{cases} \frac{1}{\beta} e^{-\frac{x}{\beta}} & , \quad x \geq 0 \\ 0 & , \quad \text{annars} \end{cases}$$

Då gäller

$$F_X(x) = 1 - e^{-\frac{x}{\beta}}, \quad \mathbb{E}[X] = \beta, \text{ samt } V(X) = \beta^2$$

13.3 Normalfördelningen

Definition 13.3. Normalfördelningen En kontinuerlig s.v., X , är normalfördelad med parametrar μ och σ , vilket skrivs $X \sim \mathcal{N}(\mu, \sigma)$, om

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Då gäller:

$$\mathbb{E}[X] = \mu, \text{ och } V(X) = \sigma^2$$

Om $X \sim \mathcal{N}(0, 1)$ skrivs $f_X(x) = \varphi(x)$ och

$$F_X(x) = \int_{-\infty}^x \varphi(t) dt = \Phi(x)$$

Tabellen på sid 391 ger $\Phi(x)$, alltså fördelningsfunktion för en $\mathcal{N}(0, 1)$ -fördelad s.v. MEN: om $X \sim \mathcal{N}(0, 1)$ så gäller

$$Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

(Steg 1: centrera, steg 2: skala. Rita bilder)

Därför kan $F_X(x)$ uttryckas med Φ , oavsett parametrar:

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = \mathbb{P}\left(Z \leq \frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

LÄXA: Visa att $\mathbb{E}[Z] = 0$ och $V(Z) = 1$, där $Z = \frac{X - \mu}{\sigma}$ och $X \sim \mathcal{N}(0, 1)$.

Exempel 13.2. Pelle och Palle går upp och ner i vikt lite slumpmässigt. Pelles vikt, X är $\mathcal{N}(72, 1)$ -fördelad och Palles vikt, Y , är $\mathcal{N}(73, 1)$ -fördelad. Vad är sannolikheten att Pelle väger mer än Palle (vid en slumpmässigt vald tidpunkt)?

Söker $\mathbb{P}(X > Y)$, där $X \sim \mathcal{N}(72, 1)$ och $Y \sim \mathcal{N}(73, 1)$. Notera: $\mathbb{P}(X > Y) = \mathbb{P}(X - Y > 0)$, så låt oss sätta $W = X - Y$ och fundera över den.

$$\mathbb{E}[W] = \mathbb{E}[X - Y] \stackrel{\text{regler}}{=} \mathbb{E}[X] - \mathbb{E}[Y] = 72 - 73 = -1$$

$$V(W) = V(X - Y) \stackrel{\text{oberoende}}{=} V(X) + V(Y) = 1^2 + 1^2 = 2$$

Så gäller $W \sim \mathcal{N}(-1, \sqrt{2})$?

Sats 13.3. Linjärkombination av $\mathcal{N}$ -fördelade s.v. Låt $X_1 \sim \mathcal{N}(\mu_1, \sigma_1), \dots, X_n \sim \mathcal{N}(\mu_n, \sigma_n)$ vara oberoende och $a_1, \dots, a_n, b \in \mathbb{R}$ konstanter. Då gäller

$$Y = \sum_{i=1}^n a_i X_i + b \sim \mathcal{N}(\mu_Y, \sigma_Y)$$

där

$$\mu_Y = \sum_{i=1}^n a_i \mu_i + b, \text{ och } \sigma_Y^2 = \sum_{i=1}^n a_i^2 \sigma_i^2$$

Tillbaka till exemplet: Vi har alltså, som vi hoppades, att $W \sim \mathcal{N}(-1, \sqrt{2})$. Så

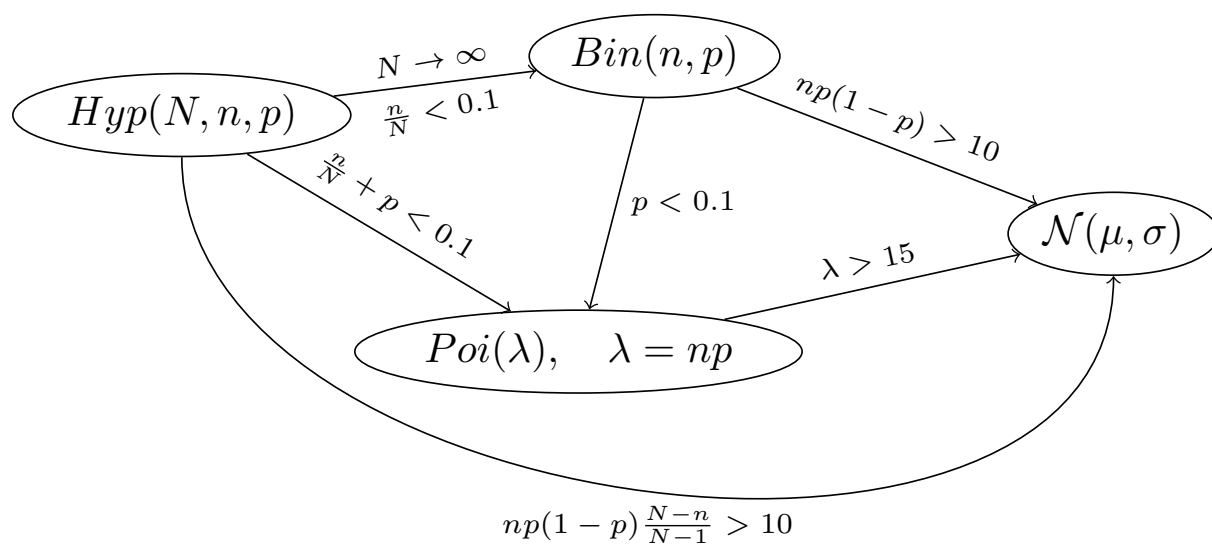
$$\begin{aligned} \mathbb{P}(W > 0) &= \mathbb{P}\left(\frac{W - \mu_W}{\sigma_W} > \frac{0 - \mu_W}{\sigma_W}\right) = \mathbb{P}\left(Z > 1/\sqrt{2}\right) = 1 - \mathbb{P}\left(Z \leq 1/\sqrt{2}\right) \\ &= 1 - \Phi\left(1/\sqrt{2}\right) = 1 - \Phi(.707) \stackrel{p.391}{=} 1 - 0.7611 = 0.2389 \end{aligned}$$

Sats 13.4. Centrala gränsvärdessatsen Låt $X_1, \dots, X_n$ vara oberoende, likafördelade s.v. med väntevärde μ och standardavvikelse σ . Då gäller

$$\mathbb{P}\left(\frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} \leq x\right) = \mathbb{P}\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq x\right) \xrightarrow{n \rightarrow \infty} \Phi(x)$$

dvs $\bar{X} \stackrel{\text{approx}}{\sim} \mathcal{N}(\mu, \sigma/\sqrt{n})$ och $n\bar{X} \stackrel{\text{approx}}{\sim} \mathcal{N}(n\mu, \sqrt{n}\sigma)$

13.4 Några approximationer



Figur 1: Några möjliga approximationer.

Föreläsning 8: 4.7

Rep.: Kontinuerliga s.v.

- Kont. likformigfördelningen: $X \sim U([a, b])$
- Exponentialfördelningen: $X \sim \text{Exp}(\lambda)$
- Normalfördelningen: $X \sim N(\mu, \sigma)$
 - Standardnormalfördelningen: $X \sim N(0, 1)$; $F_X(x) = \Phi(x)$:

$$X \sim N(\mu, \sigma) \Rightarrow Z := \frac{X - \mu}{\sigma} \sim N(0, 1)$$

- Linjärkombinationer av oberoende normalfördelade s.v.:
 $X_i \sim N(\mu_i, \sigma_i)$, $a_i, b \in \mathbb{R}$, $i = 1, \dots, n$, $Y = \sum_{i=1}^n a_i X_i + b$,

$$\Rightarrow Y \sim N\left(\sum_{i=1}^n a_i \mu_i + b, \sqrt{\sum_{i=1}^n (a_i \sigma_i)^2}\right)$$

Idag: Två-dimensionella fördelningar

- Simultan och marginella: slh. fkt. (diskret) och täthetsfkt. (kont.)
- Oberoende s.v.
- Korrelation: kovarians och korrelationskoefficient

14 Diskreta två-dim. s.v.

Kom ihåg exemplet från Förel. 2 där vi kastar två tärningar. Nu: Istället för att betrakta $Z = X + Y$ antal prickar på båda tärningar, betraktar vi (X, Y) som är en två-dim. s.v..

Exempel 14.1. Vi kastar två tärningar igen, men betraktar nu

$$\underbrace{X : \Omega \longrightarrow \{0, 1, 2\}}_{\text{antal sexor}} \quad \text{och} \quad \underbrace{Y : \Omega \longrightarrow \{0, 1, 2\}}_{\text{antal ettor}}.$$

Då gäller för sannolikheterna $\mathbb{P}(X = j \cap Y = k)$ och de "individuella" sannolikheterna $\mathbb{P}(X = j)$ och $\mathbb{P}(Y = k)$ och X och Y :

$p_{XY}(j, k) = \mathbb{P}(X = j \cap Y = k)$		k			$p_X(j) = \mathbb{P}(X = j)$
		0	1	2	
j	0	$\frac{16}{36}$	$\frac{8}{36}$	$\frac{1}{36}$	$\frac{25}{36}$
	1	$\frac{8}{36}$	$\frac{2}{36}$	0	$\frac{10}{36}$
	2	$\frac{1}{36}$	0	0	$\frac{1}{36}$
$p_Y(k) = \mathbb{P}(Y = k)$		$\frac{25}{36}$	$\frac{10}{36}$	$\frac{1}{36}$	1

Härmed kan vi också räkna ut sannolikheten för summan $Z = X + Y$ ($Z \in \{0, 1, 2, 3, 4\}$ inte samma s.v. som i Förel. 2!), t.ex.:

$$\mathbb{P}(X + Y = 1) = \mathbb{P}(X = 1 \cap Y = 0) + \mathbb{P}(X = 0 \cap Y = 1) = \frac{8}{36} + \frac{8}{36} = \frac{16}{36}$$

och på samma sätt

$$\mathbb{P}(X + Y = 0) = \frac{16}{36}, \quad \mathbb{P}(X + Y = 2) = \frac{4}{36}, \quad \text{o.s.v.}$$

eller $Z = X \cdot Y$

$$\mathbb{P}(X \cdot Y = 1) = \mathbb{P}(X = 1 \cap Y = 1) = \frac{2}{36}, \quad \mathbb{P}(X \cdot Y = 4) = 0, \quad \mathbb{P}(X \cdot Y = 0) = \frac{16 + 2 \cdot 8 + 2 \cdot 1}{36} = \frac{34}{36},$$

o.s.v..

Mer formellt:

Definition 14.1. Diskreta två-dim s.v.

- Simultan sannolikhetsfunktion:

$$p_{XY}(j, k) = \mathbb{P}(X = j \cap Y = k) = \mathbb{P}(X = j, Y = k) \geq 0, \text{ sådan att } \sum_j \sum_k p_{XY}(j, k) = 1.$$

- Marginella sannolikhetsfunktion:

$$p_X(j) = \sum_k p_{XY}(j, k), \quad p_Y(k) = \sum_j p_{XY}(j, k)$$

- Sannolikhet för händelse A , t.ex. $A = \{X = 1 \cap Y = 0\}$:

$$\mathbb{P}(A) = \sum_{(j,k) \in A} p_{XY}(j, k)$$

- Två-dim. fördelningsfunktion:

$$F_{XY}(x, y) = \mathbb{P}(X \leq x \cap Y \leq y) = \sum_{j \leq x} \sum_{k \leq y} p_{XY}(j, k) \leq 1$$

15 Kontinuerliga två-dim. s.v.

Definition 15.1. Kont. två-dim s.v.

- Simultan täthetsfunktion:

$$f_{XY}(x, y) \geq 0 \quad \forall (x, y) \in \mathbb{R}^2 \quad \text{där} \quad \int_{\mathbb{R}} \int_{\mathbb{R}} f_{XY}(x, y) dy dx = 1$$

- Marginella täthetsfunktioner:

$$f_X(x) = \int_{\mathbb{R}} f_{XY}(x, y) dy, \quad f_Y(y) = \int_{\mathbb{R}} f_{XY}(x, y) dx$$

- Sannolikhet för händelse A , t.ex. $A = \{X \leq 1 \cap Y \geq 0\}$:

$$\mathbb{P}(A) = \int_{(x,y) \in A} f_{XY}(x, y) d(x, y)$$

- Två-dim. fördelningsfunktion:

$$F_{XY}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{XY}(s, t) dt ds$$

Exempel 15.1. Givet den simultana täthetsfunktionen

$$f_{XY}(x, y) = \begin{cases} 4xy & , 0 \leq x, y \leq 1 \\ 0 & , \text{annars} \end{cases} \quad (15.1)$$

- a) Vad är de marginella tätheterna?

$$f_X(x) = \int_{\mathbb{R}} f_{XY}(x, y) dy = \int_0^1 4xy \, dy = \left[4x \frac{y^2}{2} \right]_0^1 = 2x \quad \forall x \in [0, 1], \quad 0 \text{ annars}$$

och på samma sätt

$$f_Y(y) = \int_0^1 4xy \, dx = 2y \quad \forall y \in [0, 1], \quad 0 \text{ annars}$$

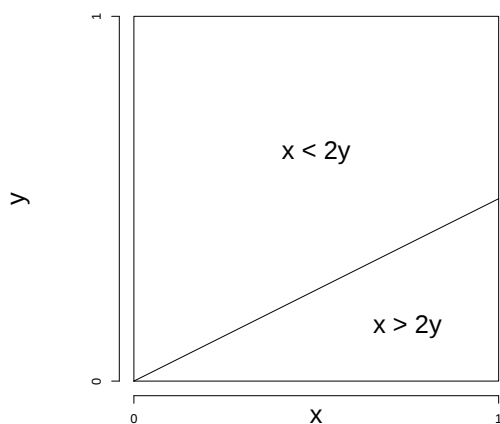
- b) Vad är sannolikheten att X är mindre än $2Y$ (d.v.s. $\mathbb{P}(X \leq 2Y)$ =? Tänk $\mathbb{P}(A)$, men vad är A)?

Vi vet att $0 \leq x, y \leq 1$ och betraktar $0 \leq x \leq 2y$. Notera: $x \leq 2y \Leftrightarrow y \geq x/2$. För alla punkter (x, y) med $y \geq x/2$ håller det. För $y \in [0, 1/2]$ har vi $x \in [0, 2y]$ med $0 \leq 2 \cdot 1/2 \leq 1$.

Därmed får vi $A = \{(x, y) \in [0, 1]^2 : y \geq x/2\}$ och:

$$\begin{aligned}\mathbb{P}(X \leq 2Y) &= \int_0^1 \int_{x/2}^1 4xy dy dx = 4 \int_0^1 x \left[\frac{y^2}{2} \right]_{x/2}^1 dx = \\ &= 4 \int_0^1 x \left(\frac{1}{2} - \frac{x^2}{8} \right) dx = \int_0^1 2x - \frac{x^2}{2} dx = \\ &= \left[x^2 - \frac{x^3}{6} \right]_0^1 = \frac{5}{6}\end{aligned}$$

OBS! Ingen uppgift som den här på tentan!



16 Oberoende stokastiska variabler

Kom ihåg oberoende händelser: $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Nu: X, Y oberoende om alla händelser definierade genom X respektive Y oberoende, dvs $\mathbb{P}(\{X \in A\} \cap \{Y \in B\}) = \mathbb{P}(\{X \in A\})\mathbb{P}(\{Y \in B\})$, $\forall A, B \subseteq \mathbb{R}$. Det visar sig att det räcker med att alla händelser av formen $\{X \leq x\}$ och $\{Y \leq y\}$ är oberoende, vilket ger

$$\underbrace{\mathbb{P}(\{X \leq x\} \cap \{Y \leq y\})}_{F_{XY}(x,y)} = \underbrace{\mathbb{P}(\{X \leq x\})}_{F_X(x)} \underbrace{\mathbb{P}(\{Y \leq y\})}_{F_Y(y)} \quad \forall x, y \in \mathbb{R}$$

Alltså: X, Y oberoende om $F_{XY}(x, y) = F_X(x)F_Y(y)$, $\forall (x, y) \in \mathbb{R}^2$. Alternativt kan man ta:

- diskret: $p_{XY}(j, k) = p_X(j)p_Y(k) \quad \forall j, k$
- kontinuerlig: $f_{XY}(x, y) = f_X(x)f_Y(y) \quad \forall x, y$

Exempel 16.1. Givet den simultana täthetsfkten in (15.1)

$$f_{XY}(x, y) = 4xy = 2x \cdot 2y = f_X(x)f_Y(Y) \quad \text{for } 0 \leq x, y \leq 1$$

(och $f_{XY}(x, y) = 0 = f_X(x)f_Y(Y)$ annars) $\Rightarrow X, Y$ oberoende.

Exempel 16.2. Vi kastar två tärningar och betraktar

$$\underbrace{X : \Omega \longrightarrow \{0, 1, 2\}}_{\text{antal sexor}} \quad \text{och} \quad \underbrace{Y : \Omega \longrightarrow \{0, 1, 2\}}_{\text{antal ettor}}.$$

Ta t.ex.

$$p_{XY}(1, 0) = \mathbb{P}(X = 1 \cap Y = 0) = \frac{8}{36} \neq \frac{10}{36} \frac{25}{36} = \mathbb{P}(X = 1)\mathbb{P}(Y = 0) = p_X(1)p_Y(0)$$

$\Rightarrow X, Y$ ej oberoende.

17 Korrelation: kovarians och korrelationskoefficient

... mått på (linjärt) beroende mellan två s.v.

Definition 17.1. Kovarians

$$C(X, Y) = \text{Cov}(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

OBS! X, Y oberoende $\Rightarrow \mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y) \Rightarrow C(X, Y) = 0 \Rightarrow X, Y$ okorreleade.
MEN $C(X, Y) = 0 \not\Rightarrow X, Y$ oberoende: det kan finnas en ej linjärt beroende!

Sats 17.1. Räknerregler för varianser (mer generellt) Låt X och Y vara godtyckliga s.v. och $a, b, c \in \mathbb{R}$ godtyckliga konstanter. Då gäller

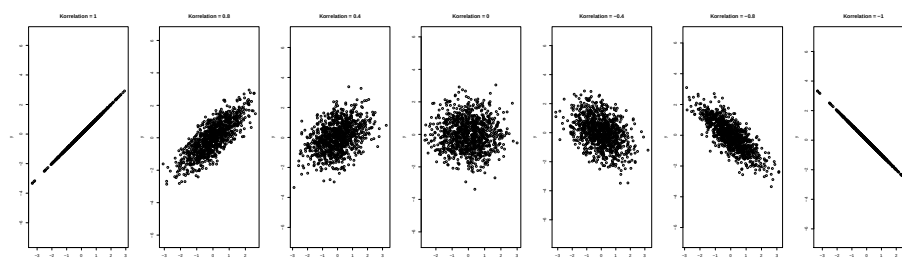
$$V(aX + bY + c) = a^2V(X) + b^2V(Y) + 2abC(X, Y)$$

Definition 17.2. Korrelationskoefficient

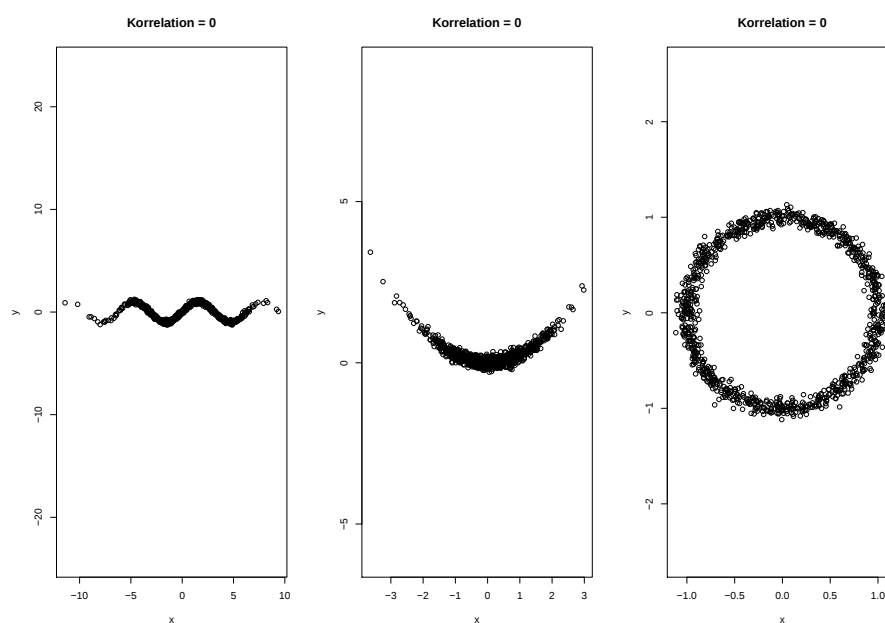
$$\rho(X, Y) = \frac{C(X, Y)}{D(X)D(Y)}$$

är en skalad variant, så att $-1 \leq \rho \leq 1$. $\rho = 0$ om X, Y okorrelerade.

NOTERA: Korrelationen är bara ett mått på *linjärt* beroende. Nedan följer exempel på korrelerade och okorrelerade slumpvariabler.



Figur 2: Exempel på korrelerade slumpvariabler.



Figur 3: Exempel på okorrelerade, men tydligt beroende, slumpvariabler.

Föreläsning 9: 2,3,5.1

18 Deskriptiv statistik

Har olika metoder för att sammanfatta och beskriva data, huvudsakligen två

- Diagram (Histogram, boxplot, cirkeldiagram, osv.)
- Numeriska sammanfattningar (Medelvärde, stickprovsvarians, korrelation, osv.)

Igenom allt detta kommer vi anta att vi har en parad datamängd, (x_i, y_i) , $i = 1, \dots, n$. Tänker på det som en mindre mängd av en (möjligen oändlig) population.

18.1 Numeriska metoder

Medelvärdet är den vanligaste beskrivande statistikan för en data mängd:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i,$$

Kan användas som en skattning av väntevärdet.

Stickprovsvariansen är ett mått på spridningen i datan, och även en skattning av variansen:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right)$$

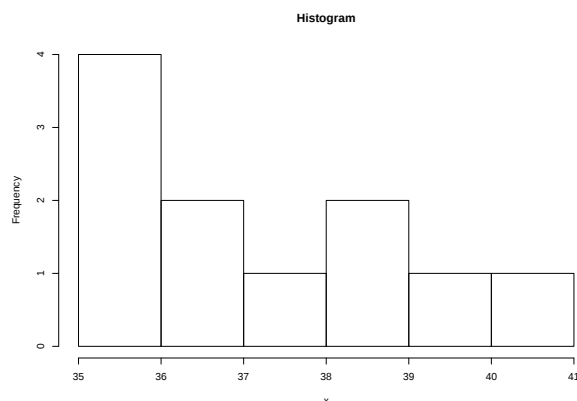
Korrelationskoefficienten är ett mått på beroendet mellan två variabler, och en skattning av korrelationen i populationen. Denna betecknas $r_{xy} = \rho(x, y)$ och ges av

$$r_{xy} = \frac{s_{xy}^2}{\sqrt{s_x^2 s_y^2}} \quad \text{där}$$

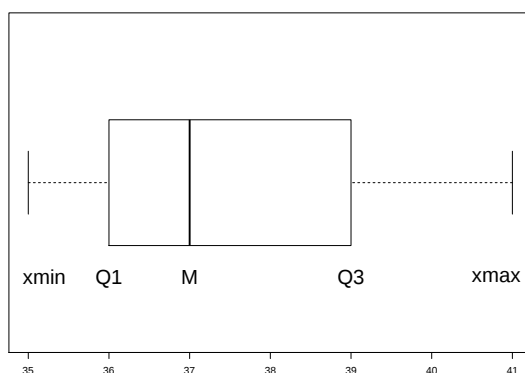
$$s_{xy}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n-1} \left(\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y} \right)$$

18.2 Diagram

Stolpdiagram eller histogram används för att illustrera fördelningen av data. Data delas upp i intervall, och andelen/antalet inom varje intervall utgör höjden på en stapel:



Box plot/lådagram illustrerar fördelningen på ett mer sammanfattande vis. En “låda” ritas mellan första och tredje kvartilerna, med ett vertikalt streck vid medianden. (Första kvartilen har 1/4 av datan till vänster, 3/4 till höger, andra 50-50 och tredje kvartilen har 3/4 till vänster och 1/4 till höger).



Ger en samlad bild av tyngdpunkt och spridning av data och används ofta för att jämföra datamängder.

18.3 Linjär regression

Antag att vi har två variabler, x och y , som vi misstänker har ett linjärt samband (dvs de är korrelerade). Ett sätt att uttrycka detta matematiskt är att vi antar

$$y = a + bx + \varepsilon$$

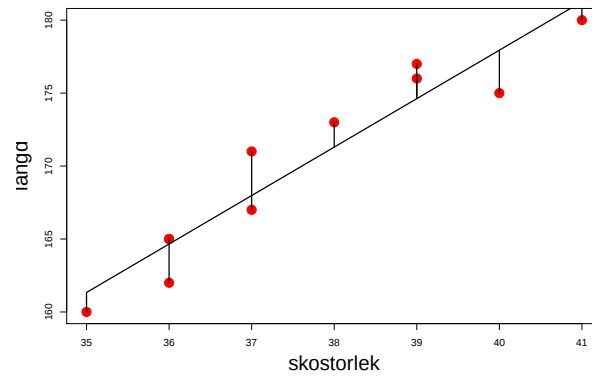
där ε är brus (dvs någon slumpvariabel med väntevärde 0). Hur kan vi skatta a och b ? Man kan visa att minsta kvadrat-metoden ger följande lösningar:

$$\hat{a} = \bar{y} - \hat{b}\bar{x} \quad \text{och} \quad \hat{b} = \frac{s_{xy}^2}{s_x^2}$$

L9MA30, LGMA30

Detta ger ett sätt att förutsäga ett y -värde för en ny observation med $x = x^*$, genom

$$y^* = a + bx^*.$$

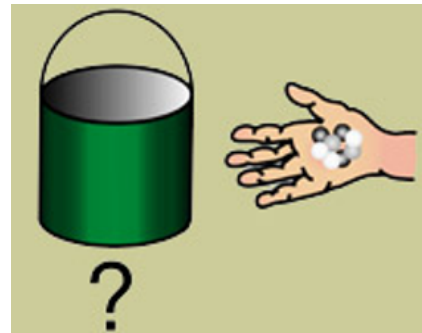


Föreläsning 10: Repetition

19 Repetition sannolikhetsdelen



(a) Sannolikhetslära.



(b) Statistik.

Figur 4: De två delarna i kursen.

Population		Stickprov (skattare)	
$F_X(x)$	Fördelningsfunktion		
$p_X(k), f_X(x)$	Täthets/sannolikhets-funktion	Histogram	
$\mathbb{E}[X]$	Väntevärde	$\bar{x}$	Medelvärde
$V[X]$	Varians	s^2	Stickprovsvarians
$C(X, Y)$	Kovarians	s_{xy}^2	
$\rho(X, Y)$	Korrelation	r_{xy}	Korrelationskoefficient

Utfall, utfallsrum, händelser. Vi gör ett försök, och vet att vi kommer observera något möjligt *utfall* (t ex antal prickar, en viss bräda). Mängden av alla utfall kallas *utfallsrum*, betecknas Ω ($\Omega = \{1, \dots, 6\}$, $\Omega = \{\text{bräda}_1, \dots, \text{bräda}_N\}$). En delmängd av utfallsrummet kallas *händelse*, A ($A = \{1, 3, 5\}$ eller $A = \text{“brädan är längre än 4 meter”}$).

19.1 Sannolikhet

a) Klassisk: Om alla utfall har samma sannolikhet gäller

$$\mathbb{P}(A) = \frac{\text{antal utfall i } A}{\text{antal möjliga utfall}} = \frac{\text{gynnsamma}}{\text{möjliga}} = \frac{|\{3, 4, 5, 6\}|}{|\{1, \dots, 10\}|} = \frac{4}{10}$$

b) Kolmogorov:

- (i) $0 \leq \mathbb{P}(A) \leq 1$,
- (ii) $\mathbb{P}(\Omega) = 1$,
- (iii) $A \cap B = \emptyset \Rightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

Räkneregler:

- (i) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
- (ii) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$

Oberoende A och B oberoende om $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

OBS: Oberoende $\neq$ disjunkta!

19.2 Betingad sannolikhet

Definition 19.1. Betingad sannolikhet: $\mathbb{P}(A \text{ givet } B) = \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$. Ger även: $\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B)$.

Sats 19.1. *Lagen om total sannolikhet:* $B_1, \dots, B_n$ disjunkta och $B_1 \cup B_2 \dots \cup B_n = \Omega \Rightarrow \mathbb{P}(A) = \mathbb{P}(A \cap B_1) + \dots \mathbb{P}(A \cap B_n) = \sum_{i=1}^n \mathbb{P}(A|B_i)\mathbb{P}(B_i)$

Notera: B_1 och $B_2 = B_1^c$ uppfyller villkoren (med $n = 2$).

Sats 19.2. *Bayes sats "Vända på betingningen"*

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(B_i \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\mathbb{P}(A)} \stackrel{LTS}{=} \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum_{i=1}^n \mathbb{P}(A|B_i)\mathbb{P}(B_i)}$$

19.3 Slumpvariabler

$X : \Omega \rightarrow \mathbb{R}$ med $\omega \in \Omega$ och $X(\omega) \in \mathbb{R}$.

Diskreta (antal brädbitar)	Kontinuerliga (brädlängd)
Sannolikhetsfunktion: $p_X(k) = \mathbb{P}(X = k)$, där $p_X(k) \geq 0$ och $\sum_k p_X(k) = 1$	Täthetsfunktion $f_X(x) \approx \frac{P(X \leq x + \Delta x) - P(X \leq x)}{\Delta x}$, $f_X(x) \geq 0$ och $\int_{\mathbb{R}} f_X(x) dx = 1$
	<u>Fördelningsfunktion</u> $F_X(x) = \mathbb{P}(X \leq x) =$
$\sum_{k \leq x} p_X(k)$	$\int_{-\infty}^x f_X(t) dt$
	<u>Väntevärde</u> $\mathbb{E}[X] =$
$\sum_k k p_X(k)$	$\int x f_X(x) dx$
	<u>Varians</u> $V(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] =$ $\mathbb{E}[X^2] - (\mathbb{E}[X])^2$
	<u>Standardavvikelse</u> $D(X) = \sqrt{V(X)}$

19.4 Några diskreta fördelningar

19.4.1 Hypergeometrisk

$X \sim \text{Hyp}(N, n, p)$, X = “antalet svarta kulor när vi drar n ur totalt N där $m = Np$ är svarta”.

$$p_X(k) = \frac{\binom{Np}{k} \binom{N(1-p)}{n-k}}{\binom{N}{n}}, \quad 0 \leq k \leq Np, \quad 0 \leq n-k \leq N(1-p)$$

$$\mathbb{E}[X] = np$$

$$V(X) = np(1-p) \frac{N-n}{N-1}$$

19.4.2 Binomial

$X \sim \text{Bin}(n, p)$, X = “antalet lyckade av n oberoende försök, där varje lyckas med sannolikhet p ”.

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, \dots, n$$

$$\mathbb{E}[X] = np$$

$$V(X) = np(1-p)$$

19.4.3 Poisson

$X \sim \text{Poi}(\lambda)$, X = “antal händelser inom ett visst tidsintervall”. Notera om $X \sim \text{Poi}(\lambda)$ är antalet händelser per tidsenhet så är antalet händelser per t tidsenheter $\text{Poi}(\lambda t)$ -fördelat.

$$p_X(k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, \dots$$

$$\mathbb{E}[X] = \lambda$$

$$V(X) = \lambda$$

19.5 Några kontinuerliga fördelningar

19.5.1 Likformig

$X \sim U([a, b])$.

$$\begin{aligned} f_X(x) &= \frac{1}{b-a}, x \in [a, b] \\ \mathbb{E}[X] &= \frac{b+a}{2} \\ V(X) &= \frac{(b-a)^2}{12} \end{aligned}$$

19.5.2 Exponential

$X \sim \exp(\beta)$, X = "tid mellan händelser i en Poisson-process" (med $\beta = 1/\lambda$).

$$\begin{aligned} f_X(x) &= \beta e^{-\beta}, x \geq 0 \\ \mathbb{E}[X] &= \beta \\ V(X) &= \beta^2 \end{aligned}$$

19.5.3 Normal

$X \sim \mathcal{N}(\mu, \sigma)$.

$$\begin{aligned} f_X(x) &= \frac{1}{\sqrt{2\pi}\sigma}, \quad x \in \mathbb{R} \\ \mathbb{E}[X] &= \mu \\ V(X) &= \sigma^2 \end{aligned}$$

20 Normalfördelningens egenskaper

Sats 20.1 (Linjärkombination av normalfördelade s.v). Låt $X_1, \dots, X_n$ vara oberoende normalfördelade med väntevärden $\mu_1, \dots, \mu_n$ och standardavvikelser $\sigma_1, \dots, \sigma_n$, och låt $a_1, \dots, a_n, b \in \mathbb{R}$. Då gäller

$$Y = \sum_i a_i X_i + b \sim \mathcal{N}(\mu_Y, \sigma_Y),$$

där $\mu_Y = \sum_i a_i \mu_i + b$ och $\sigma_Y^2 = \sum_i a_i^2 \sigma_i^2$.

Notera: detta medför att om $X \sim \mathcal{N}(\mu, \sigma)$ så gäller $\frac{X-\mu}{\sigma} =: Z \sim \mathcal{N}(0, 1)$.

Sats 20.2. Centrala gränsvärdeessatsen Låt $X_1, \dots, X_n$ vara oberoende, likafördelade s.v. med väntevärde μ och standardavvikelse σ . Då gäller

$$\mathbb{P}\left(\frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} \leq x\right) = \mathbb{P}\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq x\right) \xrightarrow{n \rightarrow \infty} \Phi(x)$$

dvs $\bar{X} \overset{approx}{\sim} \mathcal{N}(\mu, \sigma/\sqrt{n})$ och $n\bar{X} \overset{approx}{\sim} \mathcal{N}(n\mu, \sqrt{n}\sigma)$

21 Tvådimensionella slumpvariabler

21.1 Diskreta

Definition 21.1. Diskreta två-dim s.v.

- Simultan sannolikhetsfunktion:

$$p_{XY}(j, k) = \mathbb{P}(X = j \cap Y = k) = \mathbb{P}(X = j, Y = k) \geq 0, \text{ sådan att } \sum_j \sum_k p_{XY}(j, k) = 1.$$

- Marginella sannolikhetsfunktion:

$$p_X(j) = \sum_k p_{XY}(j, k), \quad p_Y(k) = \sum_j p_{XY}(j, k)$$

- Sannolikhet för händelse A , t.ex. $A = \{X = 1 \cap Y = 0\}$:

$$\mathbb{P}(A) = \sum_{(j,k) \in A} p_{XY}(j, k)$$

- Två-dim. fördelningsfunktion:

$$F_{XY}(x, y) = \mathbb{P}(X \leq x \cap Y \leq y) = \sum_{j \leq x} \sum_{k \leq y} p_{XY}(j, k) \leq 1$$

21.2 Kontinuerliga

Definition 21.2. Kont. två-dim s.v.

- Simultan täthetsfunktion:

$$f_{XY}(x, y) \geq 0 \quad \forall (x, y) \in \mathbb{R}^2 \quad \text{där} \quad \int_{\mathbb{R}} \int_{\mathbb{R}} f_{XY}(x, y) dy dx = 1$$

- Marginella täthetsfunktioner:

$$f_X(x) = \int_{\mathbb{R}} f_{XY}(x, y) dy, \quad f_Y(y) = \int_{\mathbb{R}} f_{XY}(x, y) dx$$

- Sannolikhet för händelse A , t.ex. $A = \{X \leq 1 \cap Y \geq 0\}$:

$$\mathbb{P}(A) = \int_{(x,y) \in A} f_{XY}(x,y) d(x,y)$$

- Två-dim. fördelningsfunktion:

$$F_{XY}(x,y) = \int_{-\infty}^x \int_{-\infty}^y f_{XY}(s,t) dt ds$$

21.3 Oberoende stokastiska variabler

X, Y oberoende om $F_{XY}(x,y) = F_X(x)F_Y(y)$, $\forall (x,y) \in \mathbb{R}^2$. Alternativt kan man ta:

- diskret: $p_{XY}(j,k) = p_X(j)p_Y(k) \quad \forall j,k$
- kontinuerlig: $f_{XY}(x,y) = f_X(x)f_Y(y) \quad \forall x,y$

21.4 Korrelation: kovarians och korrelationskoefficient

Notera: $\mathbb{E}[XY] = \sum_j \sum_k jk p_{XY}(j,k)$ och $\mathbb{E}[XY] = \int \int xy f_{XY}(x,y) dx dy$ i kontinuerliga fallet.

Definition 21.3. Kovarians

$$C(X,Y) = Cov(X,Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

OBS! X, Y oberoende $\Rightarrow \mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y) \Rightarrow C(X,Y) = 0 \Rightarrow X, Y$ okorrelerade.
MEN $C(X,Y) = 0 \not\Rightarrow X, Y$ oberoende: det kan finnas en ej linjärt beroende!

Definition 21.4. Korrelation

$$\rho(X,Y) = \frac{C(X,Y)}{D(X)D(Y)}$$

är en skalad variant, så att $-1 \leq \rho \leq 1$. $\rho = 0$ om X, Y okorrelerade.

Föreläsning 11: 4.6, 4.8, 5.3

Repetition: Kovarians och korrelation

Kovarians:

$$C(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

X, Y oberoende $\Rightarrow C(X, Y) = 0$ (okorrelerade), omvändningen gäller ej!

Korrelation:

$$\rho(X, Y) = \frac{C(X, Y)}{D(X)D(Y)} \in [-1, 1]$$

Regressionslinje: $y = ax + b$ Skattare:

$$\hat{a} = \bar{y} - \hat{b}\bar{x}$$

$$\hat{b} = \frac{S_{xy}^2}{S_x^2}$$

Exempel 21.1. Korrelation: låt (X, Y) vara en tvådimensionell slumpvariabel med täthetsfunktion

$$f_{XY}(x, y) = \begin{cases} 2, & 0 \leq y \leq x \leq 1 \\ 0, & \text{annars} \end{cases}$$

Beräkna $\rho(X, Y)$. Behöver $\mathbb{E}[X], \mathbb{E}[Y], \mathbb{E}[XY], V(X), V(Y)$.

$$\begin{aligned} \mathbb{E}[X] &= \int_0^1 \int_0^1 x f_{XY}(x, y) dx dy = 2 \int_0^1 x \int_0^x dy dx = 2 \int_0^1 x^2 dx = \frac{2}{3} \\ \mathbb{E}[Y] &= 2 \int_0^1 \int_1^x y dy dx = \int_0^1 x^2 dx = \frac{1}{3} \\ \mathbb{E}[XY] &= 2 \int_0^1 x \int_1^x y dy dx = \int_0^1 x^3 dx = \frac{1}{4} \\ \mathbb{E}[X^2] &= 2 \int_0^1 x^2 \int_0^x dy dx = 2 \int_0^1 x^3 dx = 1/2 \\ \mathbb{E}[Y^2] &= 2 \int_0^1 \int_0^x y^2 dy dx = 2 \int_0^1 \frac{x^3}{3} dx = \frac{1}{6}, \end{aligned}$$

vilket ger $V(X) = V(Y) = \frac{1}{18}$ och $Cov(X, Y) = \frac{1}{4} - \frac{2}{9} = \frac{1}{36}$. Slutligen

$$\rho(X, Y) = \frac{Cov(X, Y)}{D(X)D(Y)} = \frac{1/36}{1/18} = \frac{1}{2}.$$

Tolkning...?

22 Inledning till statistik

Vad är data, stickprov och statistik?

- Data är en samling fakta/mätningar ur vilken man kan erhålla information.
- En observation är en punkt i en data mängd.
- Ett stickprov är en mängd observationer från en population.
- Antalet observationer i ett stickprov kallas stickprovsstorlek.
- Statistik är
 - vetenskapen om att
 - * samla in data
 - * beskriva data (histogram, box plot, osv.)
 - * tolka data $\rightarrow$ inferens (prognos, hypotesprövning, osv)
 - Sammanfattning av data/stickprov (t ex medelvärde, osv)

22.1 Punktskattningar

22.1.1 Väntevärde

$$\hat{\mu} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

22.1.2 Andelar

Antag $X \sim \text{Bin}(n, p) \rightarrow \mathbb{E}[X] = np$.

$$\hat{p} = \frac{n\hat{\mu}}{n} = \frac{\# \text{ ettor}}{n} = \hat{\mu}$$

22.1.3 Varians

$$\hat{\sigma}^2 = s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right), \quad (\text{LÄXA})$$

eller

$$\hat{\sigma}^2 = s_*^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

22.1.4 Kovariansen

$$C(\hat{X}, Y) = S_{xy}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

22.2 Egenskaper för skattare

Varför delar vi med $n-1$ och inte med n ? Vad är en bra skattare? Man vill att $\mathbb{E}[(\hat{\tau} - \tau)^2]$ ska vara litet, alltså genomsnittliga kvadratiske felet. Det visar sig att detta kan delas upp i två delar, som ger följande önskvärda egenskaper:

- a) Väntevärdesriktig: $\mathbb{E}[\hat{\mu}] = \mu$, $\mathbb{E}[\hat{p}] = p$, $\mathbb{E}[\hat{\sigma}^2] = \sigma^2$, osv.
- b) Effektiv: $V(\hat{\mu})$ liten.

Exempel 22.1. Skillnaden mellan s^2 och s_*^2 . Antag $\mathbb{E}[X] = \mu$ och $V(X) = \sigma^2$. Notera

- $s_*^2 = \frac{n-1}{n} s^2$.
- $\mathbb{E}[X^2] = V(X) + (\mathbb{E}[X])^2$
- $\mathbb{E}[\bar{X}] = \mu$
- $V(\bar{X}) = \frac{\sigma^2}{n}$

$$\begin{aligned} \mathbb{E}[s^2] &= \mathbb{E} \left[\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \right] = \frac{1}{n-1} \left(\sum_{i=1}^n \mathbb{E}[X_i^2] - n\mathbb{E}[\bar{X}^2] \right) = \\ &= \frac{1}{n-1} (n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2) = \frac{(n-1)\sigma^2}{n-1} = \sigma^2, \end{aligned}$$

alltså är s^2 väntevärdesriktig. Men

$$\mathbb{E}[s_*^2] = \frac{n-1}{n} \mathbb{E}[s^2] = \frac{n-1}{n} \sigma^2 \neq \sigma^2,$$

alltså är s_*^2 ej väntevärdesriktig.

Å andra sidan:

$$V(s_*^2) = \left(\frac{n-1}{n} \right)^2 V(s^2) < V(s^2)$$

så s_*^2 effektivare än s^2 .

22.3 Intervallskattningar

Hittills har vi inte sagt nåt om osäkerheten i skattningarna. Finns det något alternativ? Kan vi t ex ge ett intervall $[A, B]$ sådant att $\mathbb{P}(\mu \in [A, B]) = 0.95$? Notera: μ är *okänt* men inte *slumpmässigt*. Det som är slumpmässigt måste vara intervallgränserna, eftersom dessa måste baseras på stickprovet.

Ett sätt att formulera frågan ovan är om vi kan ge ett ε sådant att

$$\mathbb{P}(|\hat{\mu} - \mu| \leq \varepsilon) = 0.95 \Leftrightarrow \mathbb{P}(-\varepsilon \leq \hat{\mu} - \mu \leq \varepsilon) \Leftrightarrow \mathbb{P}(\hat{\mu} - \varepsilon \leq \mu \leq \hat{\mu} + \varepsilon)$$

Varifrån får vi ε ? CGS! Vet att $\hat{\mu} = \bar{X} \stackrel{approx}{\sim} \mathcal{N}(\mu, \sigma/\sqrt{n}) \Rightarrow \hat{\mu} - \mu \stackrel{approx}{\sim} \mathcal{N}(0, \sigma/\sqrt{n})$. Mer om detta nästa föreläsning.

22.4 Några approximationer (revisited)

Sagt tidigare:

$$X \sim Hyp(N, n, p) \Rightarrow X \stackrel{approx}{\sim} \mathcal{N}\left(np, \sqrt{np(1-p)} \frac{N-n}{N-1}\right),$$

$$\text{om } np(1-p) \frac{N-n}{N-1} > 10$$

$$X \sim Poi(\lambda) \Rightarrow X \stackrel{approx}{\sim} \mathcal{N}(\lambda, \sqrt{\lambda})$$

$$\text{om } \lambda > 15$$

$$X \sim Bin(n, p) \Rightarrow X \stackrel{approx}{\sim} \mathcal{N}(np, \sqrt{np(1-p)})$$

$$\text{om } np(1-p) > 10$$

Varför? Låt oss titta på binomial-fallet. Låt $X \sim Bin(n, p)$, så att $\mathbb{E}[X] = np$ och $V(X) = np(1-p)$. Då kan vi skriva

$$X = \sum_{i=1}^n Y_i, \text{ där } Y_i \sim Bern(p),$$

och alla Y_i oberoende. Då gäller $\mathbb{E}[Y_i] = p$, $V(Y_i) = p(1-p)$. Vi söker

$$\begin{aligned} \mathbb{P}(X \leq x) &= \mathbb{P}\left(\sum_{i=1}^n Y_i \leq x\right) = \mathbb{P}\left(\frac{\sum_i Y_i - n\mathbb{E}[Y_i]}{\sqrt{nD(Y_i)}} \leq z\right) \\ &= \mathbb{P}\left(\frac{\sum_i Y_i - \overbrace{np}^{\mathbb{E}[X]}}{\underbrace{\sqrt{np(1-p)}}_{D(X)}} \leq z\right) \stackrel{CGS}{\approx} \Phi(z) \text{ för stort } n. \end{aligned}$$

L9MA30, LGMA30

Alltså $X \overset{approx}{\sim} \mathcal{N}(np, \sqrt{np(1-p)})$ för “stora” n , vilket beror på p .

Föreläsning 12: 5.3

Repetition: Skattare

- En bra skattare är väntevärdesriktig (t ex $\mathbb{E}[\bar{x}] = \mu$) och effektiv (liten varians).
- Med Centrala Gränsvärdesatsen får vi

$$X \sim \text{Bin}(n, p) \xrightarrow{np(1-p) > 10} X \stackrel{\text{approx}}{\sim} \mathcal{N}\left(np, \sqrt{np(1-p)}\right)$$

$$X \sim \text{Hyp}(N, n, p) \xrightarrow{np(1-p)\frac{N-n}{N-1} > 10} X \stackrel{\text{approx}}{\sim} \mathcal{N}\left(np, \sqrt{np(1-p)\frac{N-n}{N-1}}\right)$$

$$X \sim \text{Poi}(\lambda) \xrightarrow{\lambda \geq 15} X \stackrel{\text{approx}}{\sim} \mathcal{N}(\lambda, \sqrt{\lambda})$$

Idag:

23 Konfidensintervall för väntevärdet, μ

Kommer samla in ett stickprov, $X_1, \dots, X_n$, sådant att $\mathbb{E}[X_i] = \mu$, $i = 1, \dots, n$. Kan vi mha $X_1, \dots, X_n$ konstruera ett intervall, $[A, B]$ (där alltså A och B beror på stickprovet) sådant att $\mathbb{P}(\mu \in [A, B]) = 0.95$? Eller mer generellt, sådant att

$$\mathbb{P}(\mu \in [A, B]) = 1 - \alpha?$$

Ja, det kan vi, och detta är vad som kallas ett konfidensintervall (KI) med konfidensgrad/nivå $(1 - \alpha) \cdot 100\%$. Vi ska nu härleda ett sådant intervall.

Vi vill typiskt att intervallet ska vara symmetriskt, så vi kan då skriva det som att

$$|\mu - \hat{\mu}| \leq \varepsilon \Leftrightarrow -\varepsilon \leq \mu - \hat{\mu} \leq \varepsilon \Leftrightarrow \hat{\mu} - \varepsilon \leq \mu \leq \hat{\mu} + \varepsilon,$$

så $A = \hat{\mu} - \varepsilon$ och $B = \hat{\mu} + \varepsilon$. Vad är ε ? Vi använder $\hat{\mu} = \bar{X}$, så om $X_i \sim \mathcal{N}(\mu, \sigma)$ eller n stort, så gäller (åtminstone ungefärligt) $\bar{X} \sim \mathcal{N}(\mu, \sigma/\sqrt{n})$ och alltså $\hat{\mu} - \mu \sim \mathcal{N}(0, \sigma/\sqrt{n})$.

Ny notation: Låt $\lambda_\alpha = \Phi^{-1}(1 - \alpha)$, dvs det tal sådant att för $Z \sim \mathcal{N}(0, 1)$ så gäller $\mathbb{P}(Z \geq \lambda_\alpha) = \alpha$. (BILD). Då har vi att

$$\begin{aligned} \mathbb{P}\left(-\lambda_{\alpha/2} \leq \frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}} \leq \lambda_{\alpha/2}\right) &= 1 - \alpha \Leftrightarrow \\ \mathbb{P}\left(\hat{\mu} - \lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \hat{\mu} + \lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) &= 1 - \alpha \end{aligned}$$

Så! Ett $100 \cdot (1 - \alpha)\%$ KI för μ ges av

$$I_{\mu, \sigma \text{ känt}} = \left[\bar{x} \pm \lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right]$$

Tolkning: Ett intervall som konstrueras enligt formeln ovan täcker det sanna värdet på μ med sannolikhet $1 - \alpha$. Typiskt $\alpha = 0.05$.

Trubbel? Vi känner inte σ :(Vad gör vi då?

Vi skattar σ^2 med s^2 och använder den s.k. t -fördelningen. Bakgrunden till detta är att man kan visa att med $X_1, \dots, X_n$ ett stickprov från fördelning med $\mathbb{E}[X] = \mu$ och $V(X) = \sigma^2$ så gäller att

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} \sim T_{n-1},$$

dvs istället för standard normal-fördelningen får vi en T -fördelning med $n - 1$ frihetsgrader. T -fördelningarna är väldigt lika $\mathcal{N}(0, 1)$, men med "tyngre svansar". Ju större frihetsgrad, desto mer lik $\mathcal{N}(0, 1)$. Heuristisk tolkning: eftersom vi är osäkra på skattningen av σ måste vi öka bredden på KI.

Så! Ett $100 \cdot (1 - \alpha)\%$ KI för μ med σ okänt ges av

$$I_{\mu, \sigma \text{ okänt}} = \left[\bar{x} \pm t(n-1)_{\alpha/2} \frac{s}{\sqrt{n}} \right]$$

24 Hypotesprövning med konfidensintervall

Skiss över idén till hypotestest: Vi vill testa om en parameter, t ex μ är större/mindre eller inte lika med nåt visst värde. Kalla detta värde μ_0 . Vi kommer sedan göra observationer och fråga oss "hur starkt talar dessa data emot att $\mu = \mu_0$ ". Vi kan INTE prata om "sannolikheten att $\mu = \mu_0$ ", för vi har antagit att μ är en konstant, men för oss okänd. Däremot kan vi fråga oss "hur sannolikt var vårt utfall om $\mu = \mu_0$ ".

Exempel 24.1. Vi spelar tärning! Vi misstänker på förhand att nån har mixtrat med en av tärningarna, så att den tenderar att slå sexor oftare än en rättvis tärning. För att undersöka detta slår vi tärningen 10 gånger, och noterar att vi slår 6 hela 8 gånger. Skulle vi tro att $p = P(X = 6) = 1/6$?

Antag att p verkligen är $1/6$. Hur ovanligt är vårt utfall? $Y = \text{"Antalet gånger vi slår 6"} \sim \text{Bin}(10, p)$, och under nollhypotesen är $p = 1/6$. Så vad ska vi titta på? $P(Y \geq 8) \approx 2 \cdot 10^{-5}$. Så antingen är tärningen inte rättvis, eller så har vi observerat något väldigt ovanligt. Hur ovanligt ska det vara för att vi ska välja att inte tro på nollhypotesen? Detta väljer man INNAN testet, och den nivån kallas signifikansnivån. Typiskt är denna 1, 5 eller 10 % och betecknas α .

Formellt ställer man upp hypoteserna man vill testa:

$$\begin{array}{ll}\text{Nollhypotes} & H_0 : \mu = \mu_0 \\ \text{Alternativ hypotes} & H_1 \mu \neq \mu_0\end{array}$$

Om vi observerar något som talar så starkt för H_1 att sannolikheten att för det om H_0 de facto är sann är mindre än α så förkastar vi H_0 , dvs anser oss ha bevisat att H_1 stämmer. Det visar sig att om vi förkastar H_0 precis när KI inte täcker μ_0 , så är ovanstående uppfyllt. Dvs med hypoteserna som ovan, så förkastar vi H_0 om $\mu_0 \notin I_\mu$. Om vi istället testat enkelsidigt:

$$\begin{array}{ll}H_0 : \mu \geq \mu_0 \text{ vs } H_1 : \mu \leq \mu_0 \Rightarrow \text{Förkasta } H_0 \text{ om } \mu_0 > I_\mu^{\text{övre}} \\ H_0 : \mu \leq \mu_0 \text{ vs } H_1 : \mu \geq \mu_0 \Rightarrow \text{Förkasta } H_0 \text{ om } \mu_0 < I_\mu^{\text{nedre}}\end{array}$$

Föreläsning 13: 5.3

Repetition: KI för μ om

a) σ känt:

$$I_{\mu, \sigma \text{ känt}} = \left[\bar{x} \pm \lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right]$$

b) σ okänt:

$$I_{\mu, \sigma \text{ okänt}} = \left[\bar{x} \pm t(n-1)_{\alpha/2} \frac{s}{\sqrt{n}} \right]$$

I båda fallen: konfidensgraden är $1-\alpha$, och intervallen kan användas till att göra hypotestest vid signifikansnivå α .

Hypotestprövning: Vill t ex test,

$$H_0 : \mu = \mu_0 \text{ mot}$$

$$H_1 : \mu \neq \mu_0$$

vid signifikansnivå α . Dvs behöver en regel som gör att om vi följer den förkastar vi H_0 med sannolikhet α när H_0 är sann. En sådan regel är: förkasta H_0 om $\mu_0 \notin I_\mu$, annars förkasta inte.

Idag: KI för väntevärdesskillnad och stickprov-i-par.

25 Konfidensintervall för väntevärdesskillnad

Antag att vi har två stickprov från slumpvariabler med samma σ :

$$\left. \begin{array}{l} X_1, \dots, X_{n_X} \sim \mathcal{N}(\mu_X, \sigma) \\ Y_1, \dots, Y_{n_Y} \sim \mathcal{N}(\mu_Y, \sigma) \end{array} \right\} \text{ oberoende}$$

Vi undrar: är $\mu_X = \mu_Y$? Kommer använda följande:

$$\hat{\mu}_X = \bar{X} \sim \mathcal{N}\left(\mu_X, \frac{\sigma}{\sqrt{n_X}}\right), \quad \hat{\mu}_Y = \bar{Y} \sim \mathcal{N}\left(\mu_Y, \frac{\sigma}{\sqrt{n_Y}}\right)$$

åtminstone approximativt för stora n enligt CGS. Men då har vi också för $\mu_X - \mu_Y = \hat{\mu}_X - \hat{\mu}_Y = \bar{X} - \bar{Y}$:

$$\mathbb{E}[\bar{X} - \bar{Y}] = \mathbb{E}[\bar{X}] - \mathbb{E}[\bar{Y}] = \mu_X - \mu_Y$$

$$V(\bar{X} - \bar{Y}) = V(\bar{X}) + V(\bar{Y}) = \frac{\sigma^2}{n_X} + \frac{\sigma^2}{n_Y}$$

så:

$$\bar{X} - \bar{Y} \sim \mathcal{N}\left(\mu_X - \mu_Y, \sigma\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}\right) \quad (25.1)$$

$$\frac{\bar{X} - \bar{Y}}{\sigma\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}} \sim \mathcal{N}(0, 1) \quad (25.2)$$

Notera: det är precis samma mönster som när vi härledde KI för μ : vi har något, säg T (för Teststatistika), som är $\mathcal{N}(0, 1)$, så gäller $\mathbb{P}(-\lambda_{\alpha/2} \leq T \leq \lambda_{\alpha/2}) = 1 - \alpha$. Genom att "lösa ut" det vi är intresserade av ur T får vi ett KI. I det här fallet är uttrycket i (25.2), vilket ger

100(1- α)% KI för väntevärdesskillnaden, σ känt

$$I_{\mu_X - \mu_Y, \sigma \text{ känt}} = \left[(\bar{X} - \bar{Y}) \pm \lambda_{\alpha/2} \sigma \sqrt{\frac{1}{n_X} + \frac{1}{n_Y}} \right]$$

σ okänt blir lite klurigare, men här är uttrycket:

100(1- α)% KI för väntevärdesskillnaden, σ okänt

$$I_{\mu_X - \mu_Y, \sigma \text{ okänd}} = \left[(\bar{X} - \bar{Y}) \pm t(n_X + n_Y - 2)_{\alpha/2} s_p \sqrt{\frac{1}{n_X} + \frac{1}{n_Y}} \right]$$

där s_p är den "poolade" variansen:

$$s_p^2 = \frac{(n_X - 1)s_X^2 + (n_Y - 1)s_Y^2}{n_X + n_Y - 2}$$

Läxa: visa att s_p är väntevärdesriktig, dvs $\mathbb{E}[s_p] = \sigma$.

Exempel 25.1. (5.10)

Kan vi vid signifikansnivå 5% påstå att elever i årskurs 6 och 9 tittar olika mycket på TV?

Vill testa:

$$H_0 : \mu_6 = \mu_9 \Leftrightarrow \mu_6 - \mu_9 = 0$$

$$H_1 : \mu_6 \neq \mu_9 \Leftrightarrow \mu_6 - \mu_9 \neq 0$$

Låt X svara mot åk 6 och Y mot åk 9. Testar genom att göra 95% KI för $\mu_X - \mu_Y$. Vi får data som ger oss:

$$\begin{aligned} n_X &= 6, & n_Y &= 7 \\ \bar{x} &= 151.83, & \bar{y} &= 118.43 \\ s_x &= 63.39, & s_y &= 41.99 \\ s_p &= 52.80, & t(6 + 7 - 2)_{0.025} &= 2.201 \end{aligned}$$

så

$$I_{\mu_X - \mu_Y} = \left[(151.83 - 118.43) \pm 2.201 \cdot 52.80 \sqrt{1/6 + 1/7} \right] = [-31.244, 98.054]$$

Slutsats? $0 \in I_{\mu_X - \mu_Y}$, så kan inte förkasta H_0 .

26 Konfidensintervall vid stickprov-i-par

Antag nu att vi samlar data

$$\begin{array}{cccc} X_1, & X_2, & \dots & ,X_n \\ Y_1, & Y_2, & \dots & ,Y_n \end{array}$$

där X_i och X_j är oberoende och samma för Y , men där X_i och Y_i är beroende. Ex: gör två mätningar på samma objekt, individ eller föremål. Jfr: vilopuls före/efter meditation. Då vill vi utnyttja beroendet mellan X och Y ! Betrakta $Z_i = X_i - Y_i$:

$$\mathbb{E}[Z_i] = \mathbb{E}[X_i - Y_i] = \mathbb{E}[X_i] - \mathbb{E}[Y_i] = \mu_X - \mu_Y = \Delta$$

så vi får stickprov $Z_1, \dots, Z_n$ med $\mathbb{E}[Z_i] = \Delta$ och okänd standardavvikelse $\sigma_Z \Rightarrow$ skattar σ_Z med s_Z .

100(1- α)% KI för $\Delta = \mu_X - \mu_Y$ vid stickprov-i-par

$$I_\Delta = \left[\bar{z} \pm t(n-1)_{\alpha/2} \frac{s_Z}{\sqrt{n}} \right]$$

där $z_i = x_i - y_i$, $i = 1, \dots, n$.

Exempel 26.1. Tentamen 2014-01-11, uppgift 7).

Föreläsning 14: 5.3

Repetition:

Konfidensintervall:

KI för μ : Kom ihåg: CGS ger att $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \stackrel{\text{approx}}{\sim} \mathcal{N}(0, 1)$

$$\mathbb{P} \left(-\lambda_{\alpha/2} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq \lambda_{\alpha/2} \right) = 1 - \alpha$$

där $1 - \alpha$ är konfidensgraden och α är signifikansnivån. Om σ okänt: ersätt σ med s och $\mathcal{N}$ med T_{n-1} . Vi har sett KI för:

- μ med känt σ :

$$I_{\mu, \sigma \text{ känt}} = \left[\bar{X} \pm \lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right]$$

- μ med okänt σ :

$$I_{\mu, \sigma \text{ känt}} = \left[\bar{X} \pm t(n-1)_{\alpha/2} \frac{s}{\sqrt{n}} \right]$$

- $\mu_X - \mu_Y$ med okänt σ :

$$I_{\mu_X - \mu_Y, \sigma \text{ okänt}} = \left[\bar{X} - \bar{Y} \pm t(n_X + n_Y - 2)_{\alpha/2} s_p \sqrt{1/n_X + 1/n_Y} \right]$$

med σ känt: byt ut $t(n_X + n_Y - 2)_{\alpha/2}$ mot $\lambda_{\alpha/2}$ och s_p mot σ .

- $\Delta = \mathbb{E}[X - Y] = Z$, vid stickprov-i-par:

$$I_{\Delta} = \left[\bar{Z} \pm t(n-1)_{\alpha/2} \frac{s_Z}{\sqrt{n}} \right]$$

Hypotesprövning: Om vi testar:

$$H_0 : \theta = \theta_0 \text{ mot } \theta \neq \theta_0$$

förkastar vi H_0 om $\theta_0 \notin I_{\theta}$.

Om vi testar:

$$H_0 : \theta < \theta_0 \text{ mot } \theta \geq \theta_0$$

förkastar vi H_0 om $\theta_0 \leq I_{\theta}^{\text{nedre}}$.

Om vi testar:

$$H_0 : \theta > \theta_0 \text{ mot } \theta \leq \theta_0$$

förkastar vi H_0 om $\theta_0 \geq I_{\theta}^{\text{övre}}$.

(BILDER)

27 Konfidensintervall för andelar

Kom ihåg exemplet: 100 personer tillfrågades om de åt gröt till frukost och 21 svarade ja. Detta ger $\hat{p} = 0.21$, med $X =$ "antal grötätare av n tillfrågade ur hela populationen" $\Rightarrow X \sim \text{Bin}(n, p)$. Om $np(1-p) > 10$ kan vi anta att $X \sim \mathcal{N}(np, \sqrt{np(1-p)})$, i vårt fall: $n\hat{p}(1-\hat{p}) = 16.59$. Detta ger

$$\frac{n\hat{p} - np}{\sqrt{np(1-p)}} \sim \mathcal{N}(0, 1) \stackrel{\cdot \frac{1}{n}}{\Leftrightarrow} \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} \sim \mathcal{N}(0, 1)$$

Vet inte p , vilket vi behöver för att beräkna standardavvikelsen. Lös genom att ersätta med $\hat{p}$.

100(1- α)% KI för andel p i oändlig population

$$I_p = \left[\hat{p} \pm \lambda_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right]$$

95% KI för p i vårt exempel:

$$I_p = \left[.21 \pm 1.96 \sqrt{\frac{0.1659}{100}} \right] = [0.130, 0.290]$$

Antag att vi velat testa $H_0 : p \geq 0.3$ mot $H_1 : p \leq 0.3$. Slutsats? $p_0 = 0.3 > 0.290 = I_p^{\text{övre}} \Rightarrow$ förkasta H_0 .

Nu antog vi att vi frågat 100 pers ur en oändlig population, men undersökningen gjordes väl på en skola? Låt säga att sagda skola har 500 elever. Vad händer då? $X =$ "antal grötätare" $\sim \text{Hyp}(N = 500, n = 100, p)$. Vidare har vi att $\hat{p} = 0.21$, så $n\hat{p}(1-\hat{p}) \frac{N-n}{N-1} = 16.59 \frac{400}{499} = 13.3 > 10$ så vi kan normalapproximera! Alltså

$$X \stackrel{\text{approx}}{\sim} \mathcal{N} \left(np, \sqrt{np(1-p) \frac{N-n}{N-1}} \right) \Rightarrow \frac{n\hat{p} - np}{\sqrt{np(1-p) \frac{N-n}{N-1}}} \stackrel{\text{approx}}{\sim} \mathcal{N}(0, 1)$$

100(1- α)% KI för andel p i ändlig population

$$I_p = \left[\hat{p} \pm \lambda_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n} \frac{N-n}{N-1}} \right]$$

NOTERA: enda skillnaden mot oändlig pop är "ändlighetskorrektionen" $\frac{N-n}{N-1}$. I vårt exempel skulle detta ge:

$$I_p = \left[0.21 \pm 1.96 \sqrt{\frac{0.21(1-0.21)}{100} \frac{400}{499}} \right] = [0.1385, 0.2815]$$

28 Konfidensintervall för andelskillnader

Antag nu istället att vi vill jämföra andelar i två (oändliga) populationer, mha stickprov $X_1, \dots, X_{n_1} \sim \text{Bin}(n_1, p_1)$ och $Y_1, \dots, Y_{n_2} \sim \text{Bin}(n_2, p_2)$. Vi är då intresserade av $p_1 - p_2$, vilken vi kan skatta med $\hat{p}_1 - \hat{p}_2$. Notera att

$$\mathbb{E}[\hat{p}_1 - \hat{p}_2] = p_1 - p_2, \text{ och } V(\hat{p}_1 - \hat{p}_2) = V(\hat{p}_1) + V(\hat{p}_2),$$

och om $\min_i n_i p_i (1 - p_i) > 10$ så kan vi tillämpa CGS och, på samma sätt som när vi bara var intresserade av *en* andel, få att

$$\frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}_1(1 - \hat{p}_1)/n_1 + \hat{p}_2(1 - \hat{p}_2)/n_2}} \stackrel{\text{approx}}{\sim} \mathcal{N}(0, 1)$$

vilket ger oss

100(1- α)% KI för andel $p_1 - p_2$ i oändlig population

$$I_{p_1 - p_2} = \left[\hat{p}_1 - \hat{p}_2 \pm \lambda_{\alpha/2} \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}} \right]$$

Vad händer om vi har ändliga populationer? Ändlighetskorrektur som ovan!

100(1- α)% KI för andel $p_1 - p_2$ i ändlig population

$$I_{p_1 - p_2} = \left[\hat{p}_1 - \hat{p}_2 \pm \lambda_{\alpha/2} \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} \frac{N_1 - n_1}{N_1 - 1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2} \frac{N_2 - n_2}{N_2 - 1}} \right]$$

Föreläsning 15: Repetition

Vad kommer på tentan?

- 9 uppgifter totalt
- 6 på sannolikhetsdelen ≈ 20 p
- 3 på statistikdelen ≈ 10 p
- 30 p totalt, 15 p för godkänt, 24 p för VG

29 Sannolikhetsdelen

På tentan:

1. Visa någon (o)likhet för sannolikhet för händelser
2. Betingad sannolikhet, Bayes sats
3. Diskreta slumpvariabler
4. Kontinuerlig slumpvariabel och/eller bevis för $\mathbb{E}, V, C$
5. Diskret eller kontinuerlig tvådimensionell slumpvariabel

29.1 Sannolikhet för händelser

Allmänt:

$$\begin{aligned}0 &\leq \mathbb{P}(A) \leq 1 \\ \mathbb{P}(A^c) &= 1 - \mathbb{P}(A) \\ \mathbb{P}(A \cup B) &= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)\end{aligned}$$

OM oberoende: $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Viktigt: bevisteknik ,rita bilder!

29.2 Betingad sannolikhet/Bayes

Viktiga koncept: Betingad sannolikhet, lagen om total sannolikhet, Bayes sats.

Betingad sannolikhet:

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

Lagen om total sannolikhet: Låt $A, B_1, B_2, \dots, B_n$ vara händelser sådana att $B_1, \dots, B_n$ delar upp utfallsrummet, d.v.s. är alla disjunkta och utgör tillsammans hela utfallsrummet ($B_i \cap B_j = \emptyset \forall i \neq j$ och $B_1 \cup \dots \cup B_n = \Omega$). Då gäller lagen om total sannolikhet:

$$\mathbb{P}(A) = \mathbb{P}(B_1)\mathbb{P}(A|B_1) + \dots \mathbb{P}(B_n)\mathbb{P}(A|B_n)$$

Bayes sats: Under samma villkor som i lagen om total sannolikhet gäller följande för $i = 1, \dots, n$:

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(B_i)\mathbb{P}(A|B_i)}{\mathbb{P}(B_1)\mathbb{P}(A|B_1) + \dots + \mathbb{P}(B_n)\mathbb{P}(A|B_n)}$$

Oftast: $n = 2$, $B_1 = B$, $B_2 = B^c$.

29.3 Diskreta slumpvariabler

Viktigt: de olika fördelningarna vi pratat om och vad de beskriver (Bernoulli, Binomial, Hypergeometrisk, för-första-gången, Poisson, likformig). Sannolikhetsfunktion, fördelningsfunktion, väntevärde, varians.

$$\begin{aligned} p_X(k) &= \mathbb{P}(X = k) \\ F_X(x) &= \mathbb{P}(X \leq k) = \sum_{k \leq x} p_X(k) \\ E[X] &= \sum_k k p_X(k) \\ V(X) &= \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - \mu^2, \end{aligned}$$

där $\mathbb{E}[X^2] = \sum_k k^2 p_X(k)$.

29.4 Kontinuerliga slumpvariabler

Viktigt: de olika fördelningarna vi pratat om och vad de beskriver (likformig, exponential, normal). Täthetsfunktion, fördelningsfunktion, väntevärde, varians. Normalfördelningens egenskaper.

$$\begin{aligned} F_X(x) &= \mathbb{P}(X \leq k) = \int_{-\infty}^x f_X(t) dt \\ E[X] &= \int x f_X(x) dx \\ V(X) &= \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - \mu^2, \end{aligned}$$

där $\mathbb{E}[X^2] = \int x^2 f_X(x) dx$.

Linjärkombination av normalfördelade variabler: Om $X \sim \mathcal{N}(\mu_X, \sigma_X)$ och $Y \sim \mathcal{N}(\mu_Y, \sigma_Y)$, X och Y oberoende, och $a, b, c \in \mathbb{R}$ så

$$aX + bY + c \sim \mathcal{N}(a\mu_X + b\mu_Y + c, \sqrt{\sigma_X^2 + \sigma_Y^2})$$

Centrala gränsvärdessatsen: Om $X_1, \dots, X_n$ är oberoende likafördelade slumpvariabler med $\mathbb{E}[X_i] = \mu$ och $D(X) = \sigma$ så gäller

$$\sum_i X_i \stackrel{approx}{\sim} \mathcal{N}(n\mu, \sqrt{n}\sigma)$$

och

$$\bar{X} \stackrel{approx}{\sim} \mathcal{N}(\mu, \sigma/\sqrt{n})$$

29.5 $\mathbb{E}, V, C$

Kovarians:

$$C(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = \mathbb{E}[XY] - \mu_X \mu_Y$$

Räkneregler:

$$V(aX + bY + c) = a^2V(X) + b^2V(Y) + 2abC(X, Y)$$

Kom ihåg: om X, Y oberoende så är $C(X, Y) = 0$, omvändningen gäller ej.

Korrelation:

$$\rho(X, Y) = \frac{C(X, Y)}{D(X)D(Y)}$$

29.6 Tvådimensionella slumpvariabler

Simultan och marginella fördelningsfunktioner och sannolikhets-/täthetsfunktioner, oberoende variabler.

Diskreta:

$$\mathbb{P}(X = j \cap Y = k) = p_{XY}(j, k)$$

$$F_{XY}(x, y) = \mathbb{P}(X \leq x \cap Y \leq y) = \sum_{j \leq x} \sum_{k \leq y} p_{XY}(j, k)$$

$$p_X(j) = \sum_k p_{XY}(j, k)$$

$$\mathbb{P}(A) = \sum_{(j,k) \in A} p_{XY}(j, k)$$

$$\sum_j \sum_k p_{XY}(j, k) = 1$$

Kontinuerliga:

$$F_{XY}(x, y) = \mathbb{P}(X \leq x \cap Y \leq y) = \int_{t \leq x} \int_{s \leq y} f_{XY}(s, t) ds dt$$

$$f_X(x) = \int f_{XY}(x, y) dy$$

$$\mathbb{P}(A) = \int_{(x, y) \in A} f_{XY}(x, y) dx dy$$

$$\int \int f_{XY}(x, y) dx dy = 1$$

30 Statistikdelen

På tentan:

1. KI för andelar + hypotesprövning
2. KI för väntevärden + hypotesprövning
3. Punktskattare och egenskaper

30.1 Konfidensintervall + hypotesprövning

Kom ihåg de olika konfidensintervallen:

1. Ett väntevärde, σ känt
2. Ett väntevärde, σ okänt
3. Väntevärdesskillnad, σ känt
4. Väntevärdesskillnad, σ okänt
5. Väntevärdesskillnad, stickprov-i-par
6. En andel, oändlig och ändlig population
7. Andelsskillnad, oändlig och ändlig population

Känna igen situationerna när man ska använda de olika, frihetsgrader till t -fördelning, osv. Om man säger "Antag att observationerna kommer från fördelning med $\sigma = \dots$ " är σ känt, annars okänt. Känna igen stickprov-i-par.

Formulera hypoteser, när förkastar man H_0 ? Om vi testar:

$$H_0 : \theta = \theta_0 \text{ mot } \theta \neq \theta_0$$

förkastar vi H_0 om $\theta_0 \notin I_\theta$.

Om vi testar:

$$H_0 : \theta < \theta_0 \text{ mot } \theta \geq \theta_0$$

förkastar vi H_0 om $\theta_0 \leq I_\theta^{\text{nedre}}$.

Om vi testar:

$$H_0 : \theta > \theta_0 \text{ mot } \theta \leq \theta_0$$

förkastar vi H_0 om $\theta_0 \geq I_\theta^{\text{övre}}$.

30.2 Punktskattares egenskaper

En punktskattare $\hat{\theta}$ är väntevärdesriktig för θ om

$$\mathbb{E}[\hat{\theta}] = \theta$$

T ex: $\mathbb{E}[\bar{X}] = \mu$, $\mathbb{E}[s^2] = \sigma^2$.

En skattare $\hat{\theta}_*$ är effektivare än $\hat{\theta}$ om $V(\hat{\theta}_*) < V(\hat{\theta})$. T ex:

$$s^2 = \frac{1}{n-1} \sum_i (x_i - \bar{x})^2$$

$$s_*^2 = \frac{1}{n} \sum_i (x_i - \bar{x})^2$$

ger att $V(s_*^2) < V(s^2)$.

Väntevärdesriktig är viktigast, om två skattare är väntevärdesriktiga är den mest effektiva den bästa.