2. Stokastiska variabler (satunnaismuuttujat)

centralt begrepp: $(\Omega, \mathcal{F}, P)$ sannolikhetsrum

intuitivt: stokastisk variabel (= slumpvariabel) $\simeq$ vissa avbildningar $X: \Omega \to \mathbb{R}$ som definierade för elementära utfall $\omega \in \Omega$ (exakt def. senare).

enklaste fallet av s.v.

2.1. Diskreta stokastiska variabler (s.v)

motivation

Exempel 2.1 2 tärningar kastas,

utfallsrum $\Omega = \{(n_1, n_2) : n_1, n_2 \in \{1, \dots, 6\}\} \equiv \{1, \dots, 6\}^2$

definiera $X: \Omega \to \mathbb{R}$

$X((n_1, n_2)) = n_1 + n_2$ summan av poängtalen

möjliga värden $k = 2, 3, \dots, 12$

$$\{X = k\} \overset{\text{bet.}}{\equiv} \{\omega \in \Omega : X(\omega) = k\} \overset{\text{urbild}}{=} X^{-1}(\{k\}), \quad k = 2, 3, \dots, 12$$

exempelvis,

$$\{X = 4\} = \{(n_1, n_2) : X(n_1, n_2) = n_1 + n_2 = 4\}$$
$$= \{(1,3), (2,2), (3,1)\}$$

X är exempel på stokastisk variabel.

Anta $(\Omega, \mathcal{F}, P)$ sl-rum. Funktionen $X: \Omega \to \mathbb{R}$ är en diskret stokastisk variabel (s.v) på sl-rummet $(\Omega, \mathcal{F}, P)$ om

- X antar ett ändligt eller ett numrerbart oändligt antal värden $\{x_1, x_2, \dots\} \subset \mathbb{R}$ (obs utesluter att intervall $I \subset X(\Omega)$!)
- $\{X = x_k\} \equiv \{\omega \in \Omega : X(\omega) = x_k\} \in \mathcal{F} \quad \forall k = 1, 2, \dots$

(Måtteori: X mätbar avbildning)

> stokastisk variabel = satunnaismuuttuja = random variable

Beteckningar: om $X: \Omega \to \mathbb{R}$ stok. variabel, beteckna

$$\{X \le x\} \equiv \{\omega \in \Omega : X(\omega) \le x\}, \quad x \in \mathbb{R}$$

$$\{X \in B\} \equiv \{\omega \in \Omega : X(\omega) \in B\}, \quad \text{då } B \subset \mathbb{R}.$$

Exempel 2.1 (forts.) 2 tärningar kastas (igen!)

Följande avbildningar $X: \underbrace{\{1, \dots, 6\}^2}_{= \Omega} \to \mathbb{R}$ är exempel på diskreta stokastiska variabler:

$X((n_1,n_2)) = n_1+n_2$ (exepel ovan)

$X((n_1,n_2)) = n_1$

$X((n_1,n_2)) =$ antalet 6:or i (n_1,n_2) (dvs. 0, 1, 2)

$X((n_1,n_2)) \equiv 0 \quad \forall (n_1,n_2)\in\Omega$ etc etc

Problem: hur beskriva diskreta s.v. ?

Fördelningen (= jakauma) av en diskret s.v (= stokastisk variabel) $X:\Omega\to\mathbb{R}$ bestäms av (där $(\Omega,\mathcal{F},P)$ sl-rum)

frekvens sannolikheterna $\quad p_k \equiv P\{X=x_k\} \left(= P(\{\omega\in\Omega : X(\omega)=x_k\})\right).$

$k=1,2,\ldots$

Anta : $X:\Omega\to\mathbb{R}$ diskret s.v. Frekvens funktionen (= pistetodennäköisyysfunktio) av X är funktionen $f:\mathbb{R}\to\mathbb{R}$ för vilken

$$f(x) = P\{X=x\} \left(\overset{\text{def}}{=} P(\{\omega\in\Omega : X(\omega)=x\})\right), \quad x\in\mathbb{R}$$

(notera : $p_k = f(x_k) \quad \forall k$)

$\boxed{f = f_X}$ vid behov

egenskaper:

- $f(x)\geq 0 \quad \forall x\in\mathbb{R}$
- $f(x)>0$ endast om $x\in X(\Omega)=\{x_1,x_2,\ldots\}$ värdemängden av X.
- $\sum_k f(x_k) = 1.$

OBS $\quad f(x_k)=p_k,\ k=1,2\ldots$

(avser: $\sum_{k=1}^{\infty} f(x_k) = \lim_{n\to\infty}\sum_{k=1}^{n} f(x_k)$ om $X(\Omega)$ oändlig).

Fördelningsfunktionen (f.f. ; kertymäfunktio) till en s.v. $X:\Omega\to\mathbb{R}$ är funktionen $F:\mathbb{R}\to\mathbb{R}$ för vilken

$$F(x) = P(X\leq x) \left(\overset{\text{def}}{=} P(\{\omega\in\Omega : X(\omega)\leq x\})\right), \quad \text{då } x\in\mathbb{R}$$

$$= P(X^{-1}((-\infty,x]))$$ sl av urbild

$\boxed{F = F_X}$ vid behov

OBS f.f. till en s.v. $X:\Omega\to\mathbb{R}$ är ett centralt begrepp

f.f. reducerar s.v. $X:\Omega\to\mathbb{R}$ till reellvärd fkt $F:\mathbb{R}\to\mathbb{R}$

Låt $X:\Omega\to\mathbb{R}$ en diskret s.v.

Då bestämmer frekvens sl $\quad p_k = f(x_k) = P\{X=x_k\}$, $k=1,2\ldots$

fördelnings f. till X genom

$$F(x) = \sum_{k:\, x_k\leq x} p_k, \quad \text{då } x\in\mathbb{R}.$$

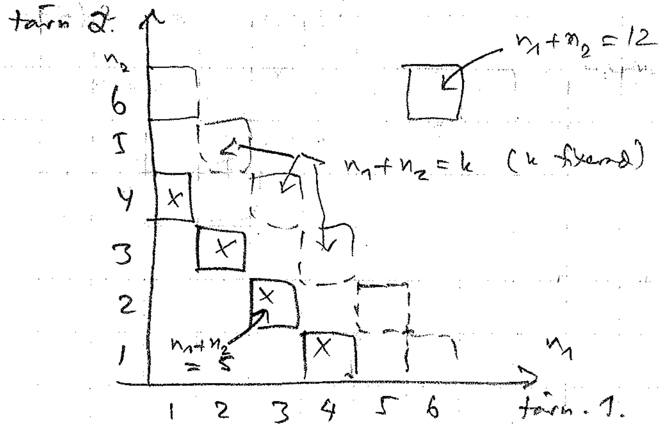
12.10

Exempel 2.2 $\quad$ 2 tärningar kastas, $\Omega=\{1,\ldots,6\}^2$ utfallsrum

$X(\omega) = n_1+n_2$, då $\omega=(n_1,n_2)\in\Omega$ $\quad$ summan av poängtalen

Bestäm frekvensfunktionen f?

Lösn. $X(\Omega) = \{2, 3, \dots, 12\} \Rightarrow f(x) = 0$ om $x \notin \{2, 3, \dots, 12\}$.



exempelvis:

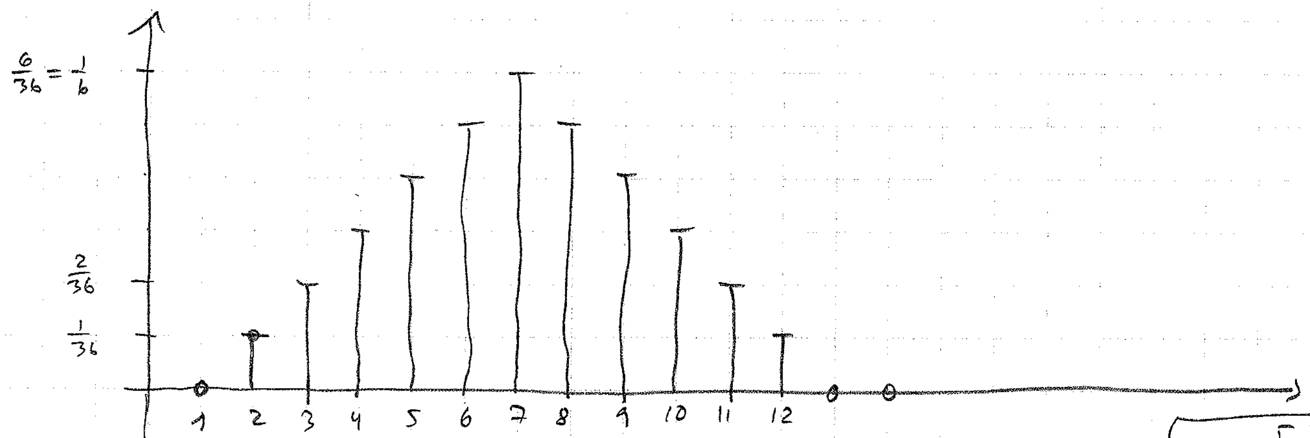
antalet element $\{\omega \in \Omega : X(\omega) = 5\} = 4$ etc $\Rightarrow$

$\{(4,1), (3,2), (2,3), (1,4)\}$

$$f(5) = P(X=5) = \frac{4}{36}$$

(antalet element i $\Omega = 36$).

som histogram (exempelvis; endast enklare fall)



Formel: $1 + \dots + n = \frac{n(n+1)}{2}$

Fördelningsfunktionen till X fås genom $F(x) = \sum_{k \le x} f(k)$,

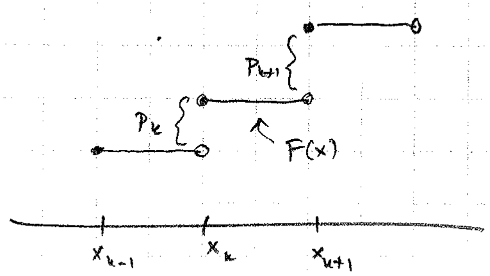
exempelvis $F(7{,}5) = \sum_{k=2}^{7} f(k) = \sum_{k=2}^{7} \left(\frac{k-1}{36}\right) = \frac{1}{36}(1+2+\dots+6) = \frac{1}{36} \cdot \frac{6 \cdot 7}{2} = \frac{7}{12}$

OBS anta $X : \Omega \to \mathbb{R}$ diskret s.v., $X(\Omega) = \{x_1, x_2, \dots\}$ där

$$\dots < x_k < x_{k+1} < x_{k+2} < \dots$$

Formeln $F(x) = \sum_{k:\, x_k \le x} p_k$, $x \in \mathbb{R}$ $\Longrightarrow$ F har följande egenskaper:

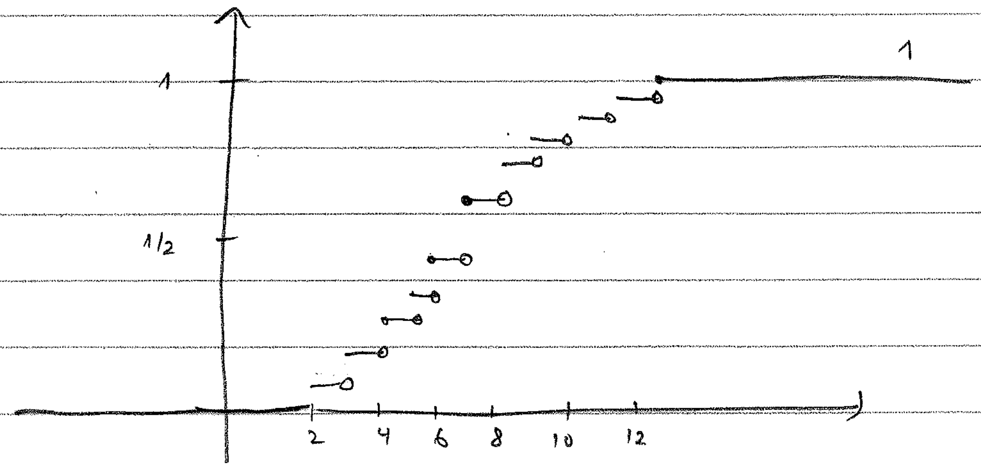
- $x \mapsto F(x)$ konstant på $[x_k, x_{k+1})$
- $x \mapsto F(x)$ växande i $\mathbb{R}$
- F har ett "hopp" med storleken $p_k > 0$ i $x = x_k$



Ex. 2.2 (fortsättning)

Fördelningsfunktion till X ?

$F(3) = f(2) + f(3) = \frac{1}{36} + \frac{2}{36} = \frac{3}{36}$ etc.



* Viktiga exempel på diskreta fördelningsfunktioner (olika tillämpningar)

A. Binomialfördelningen $(\Omega, \mathcal{F}, P)$ sl-rum

(diskreta) s.v. $X : \Omega \to \mathbb{R}$ är binomialfördelad med parameterna n, p $(n \in \mathbb{N}_+,\ 0 \le p \le 1)$ om (bet. $\mathbb{N}_+ = \{1, 2, 3, \ldots\}$)

- $X(\Omega) = \{0, 1, \ldots, n\}$ (diskret s.v.),
- $P(X = k) = P(\{\omega \in \Omega : X(\omega) = k\}) = \binom{n}{k} p^k q^{n-k}$, $k = 0, \ldots, n$

där $q = 1 - p$. $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ (binomialkoefficienterna)

Beteckna : $X \sim Bin(n, p)$

OBS: frekvensfunktionen $f_X(x) = 0$ om $x \notin \{0, \ldots, n\}$.

Typisk tolkning: binomialfördeln. beskriver sl att "lyckas" i försök med oberoende upprepningar (se kapitel 1.9):

s.v. $X \sim Bin(n,p)$ beskriver "abstrakt" version av oberoende upprepn. kan välja sl-rum $(\Omega, \mathcal{F}, P)$ etc.

Låt: A någon händelse i ett försök, $P(A) = p$ $(\Rightarrow P(A^c) = 1 - p = q)$

vid n oberoende upprepningar $\Rightarrow$

$$P(A \text{ förekommer exakt } k \text{ ggr (vid } n \text{ försök)}) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, \ldots, n$$

Frekvens ̶s̶ ̶l $p_k = P(X = k)$?

Observera:

$$(*) \quad \frac{p_{k+1}}{p_k} = \frac{\binom{n}{k+1} p^{k+1} q^{n-(k+1)}}{\binom{n}{k} p^k q^{n-k}} = \frac{p}{q} \cdot \frac{n-k}{k+1},$$

eftersom
$$\frac{\binom{n}{k+1}}{\binom{n}{k}} = \frac{\frac{n(n-1)\cdots(n-(k+1)+1)}{(k+1)!}}{\frac{n(n-1)\cdots(n-k+1)}{k!}} = \frac{n-k}{k+1}$$

ger rekursionsformel

$$\boxed{(*) \Rightarrow \quad p_{k+1} = \frac{p}{q} \cdot \frac{n-k}{k+1} \cdot p_k}$$

Anta: $X \sim Bin(5, \underbrace{0.05}_{=p}) \Rightarrow$

$$p_0 = \binom{5}{0}(0.05)^0 (0.95)^5 = (0.95)^5 \approx 0.7738$$

$$p_1 \overset{(*)}{=} \frac{p}{q} \cdot \frac{5-0}{1} p_0 = \frac{1}{19} \cdot 5 \cdot p_0 \approx 0.2036$$

$$p_2 = \frac{1}{19} \cdot \frac{4}{2} \cdot p_1 \approx 0.0214$$

$$p_3 = \frac{1}{19} \cdot \frac{3}{3} \cdot p_2 \approx 0.0011$$

$$p_4 = \frac{1}{19} \cdot \frac{2}{4} \cdot p_3 \approx 0.0003$$

$$p_5 = \frac{1}{19} \cdot \frac{1}{5} \cdot p_4 \approx 0.0000003$$

(sl för 5 "lyckade" försök av 5)

OBS $(*) \Rightarrow$ $k \mapsto p_k$ strängt växande om $\frac{p_{k+1}}{p_k} = \frac{p}{q} \cdot \frac{n-k}{k+1} > 1$

$$\Leftrightarrow \quad p(n-k) > q(k+1) \Leftrightarrow k(\underbrace{p+q}_{=1}) < np - \underbrace{q}_{1-p} = np + p - 1$$

speciellt:

$k \mapsto p_k$ strängt avtagande om $k > np + p - 1$

största frekvens ̶s̶ ̶l erhålls då $k = [np + p]$ $\equiv$ största naturliga tal m så att $m \le np + p$.

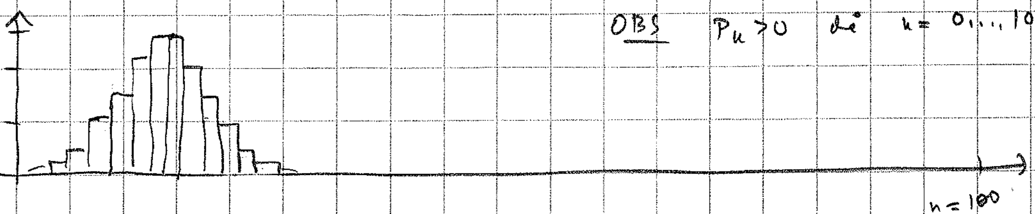
OBS Motsvarande fördelningsfunktion f.f.

$$F(x) = \sum_{k=0}^{[x]} \binom{n}{k} p^k q^{n-k} \qquad (0 \le x \le n)$$

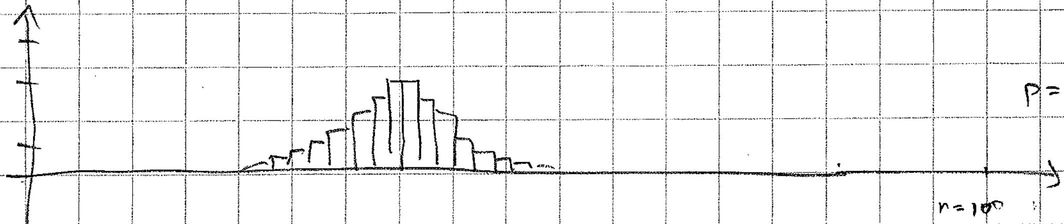
kan inte skrivas explicit. Den approximeras i de Moivre – Laplaces sats (senare 4.2)

Typiska diagram för frekvens sl till Bin (100, p):

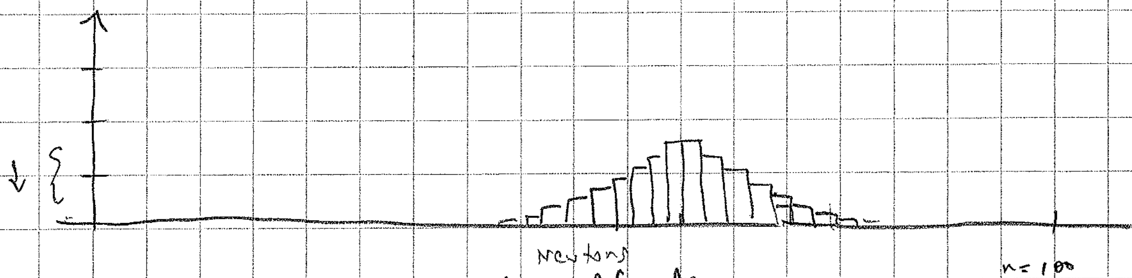
OBS $p_k > 0$ då $k = 0, \dots, 100$.



p = 0.1



p = 0.3



p = 0.5

OBS

$$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \overset{\text{Newtons binomial formeln}}{=} (\underbrace{p+q}_{=1})^n = 1$$

, dvs. uppfyller egenskapen för frekvensfunktion.

Exempel 2.3 Vid en process är felsannolikheten 0.2. 10 enheter tillverkas. Låt X = antalet defekta enheter (av 10 st).

Beräkna sl

(i) $P(X=3)$

(ii) $P(1 < X \leq 4)$

(iii) $P(X > 0)$

OBS kan även lösas med lämpliga händelser direkt från 1.9.

Lösn. Anta: tillverkningen av enheterna är oberoende av varandra

$\Rightarrow$ beskrivs av försök med oberoende upprepningar

Alltså: X diskret s.v. med $X \sim \text{Bin}(10, 0.2)$

$\boxed{n = 10, \ \text{sl } p = 0.2.}$

$P(X=3) = \binom{10}{3} (0.2)^3 (0.8)^7 \approx 0.201$ frekvens sl.

$P(1 < X \leq 4) \overset{\text{disjunkta}}{=} \sum_{j=2}^{4} f(j) = \binom{10}{2}(0.2)^2(0.8)^8 + \binom{10}{3}(0.2)^3(0.8)^7 + \binom{10}{4}(0.2)^4(0.8)^6$

$P(X=2)+P(X=3)+P(X=4)$ frekvensfkt ≈ 0.591

$\{X > 0\} = \{X = 0\}^c \Rightarrow$ komplement!

OBS $\{X > 0\} = \{$åtminstone 1 defekt enhet$\}$.

$P(X > 0) = 1 - P(X=0) = 1 - f(0) = 1 - \binom{10}{0}(0.2)^0(0.8)^{10} \approx 0.893$

13.10

B. Hypergeometrisk fördelning

$(\Omega, \mathcal{F}, P)$ sl-rum

S.v. $X: \Omega \to \mathbb{R}$ är **hypergeometriskt fördelad** med parametrarna N, K, n (där $N, K, n \in \mathbb{N}_+$, $n \le N$, $K \le N$) om

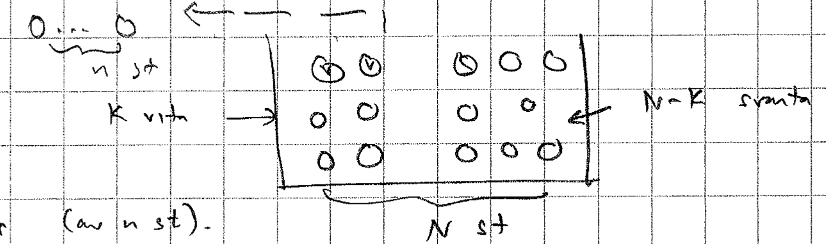
- $X(\Omega) = \{0, \dots, n\}$,
- $p_k = P(X = k) = \dfrac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$ $\quad k = 0, \dots, n$

$\boxed{\text{OBS frekvens-fkt} \quad f(x) = 0 \text{ om } x \notin \{0, \dots, n\}}$

Beteckn: $X \sim \text{Hyperg}(N, K, n)$

Tolkning: X = antalet vita kulor i dragning utan återläggning

drag n kulor utan återläggning
(se 1.6B)



$P(X = k)$ = sl få k vita kulor (av n st).

4/2/2015

C. Geometrisk fördelning

$(\Omega, \mathcal{F}, P)$ sl-rum

Diskreta

S.V. $X: \Omega \to \mathbb{R}$ är **geometriskt fördelad** med parametern p (där $0 < p < 1$) om

- $X(\Omega) = \mathbb{N}$ $\left(= \{0, 1, 2, \dots\}\right)$, uppräkneligt oändligt antal värden
- $p_k = P(X = k) = pq^k$, $k = 0, 1, \dots$ där $q = 1 - p$

Beteckna: $X \sim \text{Geom}(p)$

Tolkning: given händelse A i försök med sl $P(A) = p$, $P(A^c) = 1 - p = q$

anta godtyckligt många oberoende upprepningar av försöket

Studera första gången händelsen A uppträder

Låt X = antalet händelser A^c (="misslyckade") före 1. gången A uppträder

$\boxed{\text{intuitivt:}}$ $\quad \underbrace{A^c A^c \dots A^c}_{k \text{ st}} \overset{k+1:\text{a gången}}{A}$

$$P(\underbrace{A^c A^c \dots A^c}_{k \text{ st}} A) \overset{\text{oberoende}}{=} \underbrace{P(A^c)}_{q = 1-p}{}^k \underbrace{P(A)}_{p} = pq^k$$

OBS: inget utfallsrum beskrivet

$$\Rightarrow \quad P(X=k) = pq^k, \; k=0,\dots \quad \Rightarrow$$

$X \sim$ Geom(p). geometriskt fördelad

OBS $\quad p_k = P(X=k) = pq^k > 0 \quad \forall k = 0,1,2,\dots$, och

$$\sum_{k=0}^{\infty} p_k = \sum_{k=0}^{\infty} pq^k = p \sum_{k=0}^{\infty} q^k \underset{\text{geom. serie}}{=} p \cdot \frac{1}{\underbrace{1-q}_{=p}} = 1.$$

dvs. uppfyller villkoren för frekvenssannolikheter / frekvensfunktion.

Fördelningsfunktion:

$$F(k) = \sum_{j=0}^{k} f(j) = \sum_{j=0}^{k} pq^j = p \sum_{j=0}^{k} q^j \overset{\text{geom. serie}}{=} p \cdot \frac{1-q^{k+1}}{\underbrace{1-q}_{=p}} = 1 - q^{k+1}$$

inte 2015

Exempel 5 En tärning kastas upprepade gånger. Bestäm så att 1. sexan erhålls inom k första kasten

Exempel 2.4 se 59A

Lösn $\quad P(\text{sexa}) = \frac{1}{6} \equiv p, \quad q = 1-p = \frac{5}{6}$

tolkning: högst $k-1$ misslyckanden

$$P(\text{1. sexan bland } k \text{ första kasten}) = P(X \le k-1) = F(k) = 1 - \left(\tfrac{5}{6}\right)^k$$

Alternativ försök med upprepningar (kapitel 1.9)

$$P(\text{1. sexan bland } k \text{ första kasten}) = P(\text{åtminstone 1 sexa på } k \text{ kast})$$

$$\underset{\text{komplement}}{=} 1 - P(\text{ingen sexa på } k \text{ kast}) \overset{\text{ober.}}{=} 1 - \left(\tfrac{5}{6}\right)^k.$$

$\square$.

Exempel 2.5 (Huygens problem)

A och B kastar turvis en (symmetrisk) tärning, så att A börjar. Den person vinner som erhåller 1. sexan. Bestäm så

$$P(A \text{ vinner}) \text{ och } P(B \text{ vinner}).$$

Lösn: beskriv kasten i ordningen

$$\underset{1.}{A}, \underset{2.}{B}, \underset{3.}{A}, B, \dots$$

diskret s.v. X : $\quad X =$ antalet misslyckade kast före 1. sexan

Exempel 2.4 Anta att en apparat har sl 0.05 att få fel varje gång den används, oberoende av tidigare användningar (dvs. en användning försämrar inte apparaten). (59A)

Beräkna sl att apparaten fungerar åtminstone 20 gånger.

Lösn Låt

$X =$ # användningar av apparaten före första felet

Då är $X \sim \text{Geom}(0.05)$ med $p = 0.05$, $q = 1 - p = 0.95$, så att

$$P(X \geq 20) = P(\{X \leq 19\}^c) = 1 - P(X \leq 19)$$

$$\underset{\text{disjunkta}}{=} 1 - \sum_{k=0}^{19} P(X = k) = 1 - \sum_{k=0}^{19} p q^k$$

$$= 1 - p \sum_{k=0}^{19} q^k = 1 - p \cdot \frac{1 - q^{20}}{\underbrace{1 - q}_{= p}} = q^{20} = (0.95)^{20}$$

$$\approx 0.368$$

Påminn $\sum_{k=0}^{n} q^k = \frac{1 - q^{n+1}}{1 - q}$ då $q \neq 1$ och $n \in \mathbb{N}$ (delsumma av geometrisk serie)

inte 2015

extra fråga Beräkna sl att apparaten fungerar åtminstone 30 ggr, om apparaten redan har fungerat 10 ggr.

Lösn antagandet ("avsaknande av minne") $\Rightarrow$

$$P(X \geq 30 \mid X \geq 10) \overset{!}{=} P(X \geq 20) \approx 0.368 \quad \text{(ovan)}.$$

$$\{X = k\} = \begin{cases} k \text{ kast med poängtal } \in \{1,\dots,5\} \\ (k+1)\text{:a kastet ger poängtalet } 6. \end{cases}$$

$\boxed{P(\text{sexa}) = \frac{1}{6} \quad q = 1-p = \frac{5}{6}}$

$\Rightarrow$ geometrisk fördelning $\quad P(X=k) = \frac{1}{6}\cdot\left(\frac{5}{6}\right)^k$

sl $\quad P(A \text{ vinner}) \underset{\substack{\uparrow \\ \text{A har udda kast } 1,3,5,\dots}}{=} P(\underbrace{X = 0,2,4,\dots}_{\text{OBS räknar antalet misslyckanden!}}) = P\left(\bigcup_{k=0}^{\infty}\{X = 2k\}\right)$

P sl-fkt

$$\underset{\substack{\uparrow \\ \{X=2k+1\}\cap\{X=2l+1\}=\emptyset \\ \text{om } k\neq l}}{=} \sum_{k=0}^{\infty} P(X=2k) = \sum_{k=0}^{\infty}\frac{1}{6}\cdot\left(\frac{5}{6}\right)^{2k} = \frac{1}{6}\sum_{k=0}^{\infty}\left(\frac{25}{36}\right)^k$$

$$\underset{\text{serie}}{\overset{\text{geometrisk}}{=}} \frac{1}{6}\cdot\frac{1}{1-\frac{25}{36}} = \frac{36}{6}\cdot\frac{1}{11} = \frac{6}{11}$$

$$P(B \text{ vinner}) = P(\{A \text{ vinner}\}^c) = 1 - P(A \text{ vinner}) = 1 - \frac{6}{11} = \frac{5}{11}$$

alternativ:

$$P(B \text{ vinner}) = P\left(\bigcup_{k=0}^{\infty}\{X = 2k+1\}\right) \overset{\text{disjunkta}}{=} \sum_{k=0}^{\infty} P(X = 2k+1) = \frac{1}{6}\sum_{k=1}^{\infty}\left(\frac{5}{6}\right)^{2k-1}$$

$$\underset{\text{serie}}{\overset{\text{geom.}}{=}} \dots = \frac{5}{11}$$

$\boxed{\text{Påminn: geom serie } \sum_{k=0}^{\infty} t^k = \frac{1}{1-t} \text{ då } 0 \le t < 1}$

OBS likadant problem: myntsingling till första kronan etc.

<u>oddset</u> till A:s fördel är $\quad \frac{P(A)}{P(B)} = \frac{\frac{6}{11}}{\frac{5}{11}} = \frac{6}{5}$

<u>OBS</u> i lösningen har <u>inte</u> sl-rummet $(\Omega, \mathcal{F}, P)$ beskrivits explicit:
man <u>kan</u> välja $\quad \Omega = \{0,1,2,\dots\}, \quad P(\{k\}) = \frac{1}{6}\cdot\left(\frac{5}{6}\right)^k \overset{\text{kontrollera!}}{\Longrightarrow} \sum_{k=0}^{\infty} P(\{k\}) = 1.$

D. Poisson - fördelningen

Låt $\lambda > 0$ vara fixerad och

$$p_k = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0,1,2,\dots$$

$$\Rightarrow \sum_{k=0}^{\infty} p_k = \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\underbrace{\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}}_{= e^{\lambda}} = 1$$

$= e^{\lambda}$, Taylor serie till e^x (Analys II)

$\Rightarrow \quad p_k > 0, \; k = 0,\dots \quad$ uppfyller villkoren för frekvens sl:r

$(\Omega, \mathcal{F}, P)$ sl-rum. Diskreta s.v. $X: \Omega \to \mathbb{R}$ är **Poisson-fördelad** med parameter $\lambda > 0$ om

- $X(\Omega) = \{0, 1, 2, \ldots\}$,
- $P(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$, $k = 0, 1, 2, \ldots$

Beteckna: $X \sim \text{Poisson}(\lambda)$.

Obs: Poisson(λ) approximerar Bin(n, p) med $np \approx \lambda$, med n stort, $p \sim 0$ litet

OBS: Poisson-fördelningen Poisson(λ) erhålls som "gränsvärde" av binomialfördelningarna $X_n \sim \text{Bin}(n, p_n)$, då

$$n \to \infty \quad \text{så att} \quad np_n \xrightarrow[n\to\infty]{} \lambda$$

Orsak: Låt f_n = frekvensfunktionen till X_n,

$\lambda_n = np_n$, $n = 1, 2, \ldots$
$\Rightarrow p_n = \frac{\lambda_n}{n}$

$$\underbrace{f_n(k)}_{\text{binomial}} = \binom{n}{k} p_n^k q_n^{n-k} = \frac{(n)_k}{k!} \left(\frac{\lambda_n}{n}\right)^k \left(1 - \frac{\lambda_n}{n}\right)^{n-k} =$$

Beteckna $(n)_k = n(n-1)\cdots(n-k+1)$

Varning: ingen uppskattning av felet vid approximering

$$= \underbrace{\frac{\lambda_n^k}{k!}}_{\xrightarrow[n\to\infty]{} \frac{\lambda^k}{k!}} \cdot \underbrace{\left(1 - \frac{\lambda_n}{n}\right)^n}_{\to e^{-\lambda}} \underbrace{\frac{(n)_k}{n^k}}_{\xrightarrow[n\to\infty]{} 1} \underbrace{\left(1 - \frac{\lambda_n}{n}\right)^{-k}}_{\xrightarrow[n\to\infty]{} 1} \xrightarrow[n\to\infty]{} \frac{\lambda^k}{k!} e^{-\lambda}$$

(orsak: $\left(1 - \frac{\lambda}{n}\right)^n \to e^{-\lambda}$, $x \mapsto e^{-x}$ kontinuerlig).

$$\frac{(n)_k}{n^k} = \frac{n(n-1)\cdots(n-k+1)}{n \cdot n \cdots n} = 1 \cdot \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{k-1}{n}\right) \xrightarrow[n\to\infty]{} 1$$

(k fixerat)

Typisk tillämpning av Poisson-fördelningen

Exempel 2.6 I medeltal 48 kunder / t. anländer till en butik (oberoende av varandra). Uppskatta sl att högst 2 kunder anländer mellan kl. 12.00 och 12.05

Lösn. X = antalet kunder, egentligen $X \sim \text{Bin}\left(48, \frac{5}{60}\right)$ (5 min av 60 min)

Kan anta: $X \sim \text{Poisson}(\lambda)$, där

$$\lambda = \frac{5}{60} \cdot 48 = 4,$$

$n = 48$
$p = \frac{5}{60}$ proportion av 60 min.

Obs. ursprungliga sl "approximeras" via Poisson s.v

Exempel 7.7 Anta att antalet tryckfel på en viss sida av en lärobok i matematik följer Poisson-fördelningen med $\lambda = \frac{1}{2}$. Vilken är sannolikheten att sidan har åtminstone ett tryckfel?

(62A)

Lösn Låt $X = \#$ tryckfel på sidan. Antagandet: $X \sim \text{Poisson}(\frac{1}{2})$

Då är

$$P(X \geq 1) = P(\{X=0\}^c) = 1 - P(X=0)$$

$$= 1 - e^{-\frac{1}{2}} \frac{(\frac{1}{2})^0}{0!} = 1 - e^{-1/2} \approx 0.393$$

$\square$

$\Rightarrow$ söker (disjunkta)

sl $P(X \le 2) = P(X=0) + P(X=1) + P(X=2)$

$$= \frac{4^0}{0!}e^{-4} + \frac{4^1}{1!}e^{-4} + \frac{4^2}{2!}e^{-4} \approx 0.0183 + 0.0733 + 0.1465$$

$$= 0.2381 \quad \square.$$

(Inte 2015)

Exempel 8 Anta radioaktivt material med stor halveringstid (N år, N stort)

X = antalet atomer som sönderfaller under 10 s.

$\begin{cases} \text{varje atom sönderfaller med sl } p \text{ under } 10 \text{ s.}, \quad p \sim 0 \\ n \text{ st atomer}, \quad n \text{ stort} \end{cases}$

Exempel 2.7 se sid 62A

anta (exempelvis): $np = 4$

$X \sim Bin(n, p) \approx$ Poisson(4) (sid 61) $\Rightarrow$

sl exakt 6 atomer sönderfaller (under 10 s) =

$$P_6 = e^{-4}\frac{4^6}{6!} \approx 0.10$$

$\square$

18.10

(+) 2.2 Väntevärdet för diskreta stokastiska variabler

Anta: $(\Omega, \mathcal{F}, P)$ sl-rum

$X : \Omega \to \mathbb{R}$ diskret s.v. med

- $X(\Omega) = \{x_1, x_2, \ldots\}$ (ändlig / numrerbart oändlig)
- frekvens sl $p_k \equiv P(X = x_k) \left(\stackrel{\text{def}}{\equiv} P(\{\omega \in \Omega : X(\omega) = x_k\}) \right)$, $k = 1, 2, \ldots$

X har väntevärdet (fi. odotusarvo, eng. expected value)

$$E(X) = \sum_k x_k \cdot p_k$$

förutsatt att $\sum_k |x_k| \cdot p_k < \infty$ (Analys II: serien $\sum_k x_k \cdot p_k$ konvergerar absolut).

(X saknar väntevärde om $\sum_k |x_k| p_k = \infty$)

intuitivt: Väntevärdet för en s.v. X är ett sorts medeltal för de värden X antar (förväntat medeltal)

Exempel 2.8 anta $X(\Omega) = \{x_1, \dots, x_n\}$ (ändligt många) med jämn fördelning,

symmetriskt $p_k = P(X = x_k) = \frac{1}{n}$, $k = 1, \dots, n$

Då

$$E(X) = \sum_{k=1}^{n} x_k \cdot \underbrace{p_k}_{\frac{1}{n}} = \frac{1}{n} \sum_{k=1}^{n} x_k$$

det aritmetiska medeltalet av $x_1, \dots, x_n \in \mathbb{R}$

illustration: Tärningskast: X = poängtalet av ett kast $\in \{1, \dots, 6\} \Rightarrow$

$$E(X) = \frac{1}{6}(1+2+\dots+6) = \frac{1}{6} \cdot \frac{6 \cdot 7}{2} = \frac{7}{2}.$$

(notera: $E(X) \notin X(\Omega)$, inte värde som antas!)

Sats 2.9 Låt $X: \Omega \to \mathbb{R}$ diskret s.v., $(\Omega, \mathcal{F}, P)$ sl-rum.

(i) $X \sim \text{Bin}(n,p) \Rightarrow E(X) = np$

(ii) $X \sim \text{Hyperg}(N, K, n) \Rightarrow E(X) = n \cdot \frac{K}{N}$

(iii) $X \sim \text{Geom}(p) \Rightarrow E(X) = \frac{q}{p}$ (där $q = 1-p$)

(iv) $X \sim \text{Poisson}(\lambda) \Rightarrow E(X) = \lambda$.

Bevis (i) och (ii) senare (behöver mera teori)

(iii) $X \sim \text{Geom}(p) \Rightarrow p_k = pq^k$, $k = 0, 1, \dots$:

$$E(X) = \sum_{k=0}^{\infty} k \cdot p_k = \sum_{k=0}^{\infty} k p q^k = pq \sum_{k=1}^{\infty} k \cdot q^{k-1}$$

Kom ihåg: (Analys II, potensserier)

geometrisk serie $\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$, då $|x| < 1$

Derivera termvis (då $|x| < 1$) $\Rightarrow$

$$(*) \quad D\left(\frac{1}{1-x}\right) = \frac{1}{(1-x)^2} = D\left(\sum_{k=0}^{\infty} x^k\right) = \sum_{k=0}^{\infty} D(x^k) = \sum_{k=1}^{\infty} k x^{k-1}$$

alltså:

$$E(X) = pq \sum_{k=1}^{\infty} k q^{k-1} \overset{(*)}{=} pq \cdot \frac{1}{(\underbrace{1-q}_{p})^2} = \frac{pq}{p^2} = \frac{q}{p}.$$

(iv) $X \sim \text{Poisson}(\lambda) \Rightarrow p_k = e^{-\lambda} \frac{\lambda^k}{k!}$, $k = 0, 1, \dots \Rightarrow$

$$E(X) = \sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{k}{1\cdots(k-1)k} \lambda^k = \lambda e^{-\lambda} \underbrace{\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}}_{\text{Taylorserie } e^{\lambda}} = \lambda. \quad \square$$

<u>Exempel 2.10</u> En tärning kastas upprepade gånger. [2.4]

$X =$ antalet sexor i 100 kast

$Y =$ antalet kast före 1. sexan.

Bestäm väntevärdena $E(X)$, $E(Y)$.

<u>Lösn.</u> oberoende upprepade försök $\Rightarrow$

$$X \sim Bin(100, \tfrac{1}{6}) \overset{(i)}{\Rightarrow} \quad \boxed{\begin{matrix} n = 100 \\ p = \frac{1}{6} \end{matrix}}$$

$$E(X) = n \cdot p = \frac{100}{6} \quad (\approx 16,666\ldots)$$

kapitel 2.1 $\Rightarrow$ $Y \sim Geom(\tfrac{1}{6}) \Rightarrow$ (ty $p = \frac{1}{6}$ för sexa, $q = 1-p = \frac{5}{6}$)

$$E(Y) \overset{(iii)}{=} \frac{q}{p} = \frac{5/6}{1/6} = 5. \quad \square$$

<u>Historisk tolkning av väntevärde</u> (1600-talet)

om X (diskret) stok.v. som anger en spelares vinst under ett spel $\Rightarrow$

väntevärdet $E(X)$ anger hur mycket det lönar sig för spelaren att satsa (= betala) för att fortfarande vara på vinst

<u>St. Petersburg-paradoxen</u> (Bernoulli $\approx$ 1730)

ett symm. mynt singlas tills 1. kronan erhålls

1. kronan på n:te kastet $\Rightarrow$ vinst 2^n enheter (ofta större $n \Rightarrow$ större vinst!)

$\Rightarrow$ stok. variabel $X(k) = 2^k$, $k = 0, 1, 2\ldots$ (= spelarens vinst vid n:te kastet)

$E(X) = ?$ (= hur mycket lönar sig att satsa ?)

$$P(X = 2^n) = P(\text{1. krona på } n\text{:te kastet}) = P(\underbrace{\text{klave, klave}, \ldots, \text{klave}}_{n-1}, \underbrace{\text{krona}}_{n\text{:te}}) = \left(\frac{1}{2}\right)^{n-1} \cdot \frac{1}{2} = \frac{1}{2^n},$$

$n = 1, 2 \ldots \Rightarrow$

$$E(X) = \sum_{n=1}^{\infty} 2^n \cdot \underbrace{P(X = 2^n)}_{\frac{1}{2^n}} = 2 \cdot \frac{1}{2} + 4 \cdot \frac{1}{2^2} + \ldots + 2^n \cdot \frac{1}{2^n} + \ldots \overset{\text{divergent!}}{=} \infty \;!$$

(X saknar väntevärde). $\square$

19.10

* 2.3. Kontinuerliga fördelningsfunktioner

Anta: $(\Omega, \mathcal{F}, P)$ sl-rum. Avbildningen $X: \Omega \to \mathbb{R}$

är en (reellvärd) stokastisk variabel om

$$\{X \le a\} = \{w \in \Omega : X(w) \le a\} \in \mathcal{F} \quad \text{för alla } a \in \mathbb{R}$$

$$= X^{-1}((-\infty, a]) \text{ urbild}$$

Obs måttteori
X mätbar avbildning

OBS X stokastisk variabel $\Rightarrow$ också

$$\{b \le X \le a\} = \{w \in \Omega \mid b \le X(w) \le a\} \in \mathcal{F} \quad \forall b < a.$$

orsak

$$\{w \in \Omega : b \le X(w) \le a\} = \bigcap_{n=1}^{\infty} \{b - \tfrac{1}{n} < X(w) \le a\}$$

$$= \underbrace{\bigcap_{n=1}^{\infty} \Big[\underbrace{\{w : X(w) \le b - \tfrac{1}{n}\}^c \cap \{w : X(w) \le a\}}_{\in \mathcal{F} \quad \sigma\text{-alg.}} \Big]}_{\in \mathcal{F} \quad \text{enligt } 1.3.}$$

orsak: $b - \frac{1}{n} < X(w) \le a \quad \forall n \iff b \le X(w) \le a$

Följande typ av fördelningsfunktioner utgör "motsatsen" till diskreta fördelningar
Allmän teori för s.v. senare (kap. 2.5).

Stokastiska variabel $X: \Omega \to \mathbb{R}$ har en kontinuerlig fördelningsfunktion med frekvensfunktionen (frekv. f.) f, om

$$P(a \le X \le b) = \int_a^b f(x)\,dx \quad \forall a < b, \quad a, b \in \mathbb{R}$$

$$\| $$

$$P(\{w : a \le X(w) \le b\})$$

OBS ① ovan $\Rightarrow$
$\{w : a \le X(w) \le b\} \in \mathcal{F}$

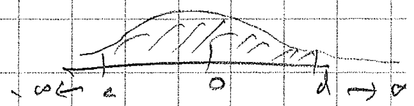
OBS ② funktionen $f: \mathbb{R} \to \mathbb{R}$ är en frekvensfunktion (för någon kont. fördelningsfunktion) om

- $f(x) \ge 0 \quad \forall x$
- f integrerbar över $\mathbb{R}$ och $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx = 1.$

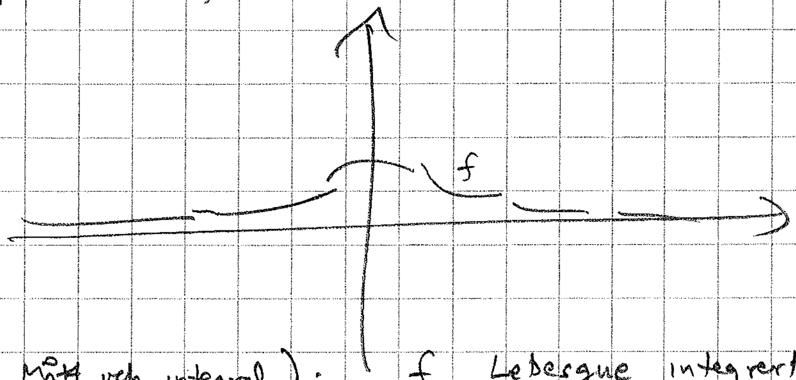
f integrerbar över $\mathbb{R}$ betyder (följande tillräckligt för kursen)

$\int_{-\infty}^{\infty} f(x)\,dx$ existerar som <u>oegentlig</u> Riemann-integral,

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{c\to-\infty}\int_c^0 f(x)\,dx + \lim_{d\to\infty}\int_0^d f(x)\,dx ,$$



<u>och</u> f (åtminstone) bitvis kontinuerlig.



<u>Allmännare</u> (Mått och integral): f Lebesgue integrerbar tillräckligt

<u>Sats 2.11</u> s.v. $X: \Omega \to \mathbb{R}$ har kontinuerlig fördelningsfunktion med frekvensf. f $\Longrightarrow$ fördelningsfunktionen $F = F_X$ är

$$F(x) = F_X(x) \overset{\text{def}}{=} P(X \le x) = \int_{-\infty}^{x} f(t)\,dt, \qquad x \in \mathbb{R}.$$

<u>Omvänt:</u>

$f(x) = F'(x)$ (derivata) för varje $x \in \mathbb{R}$ där f kontinuerlig.

"Bevis:" "minnes regel"

$$P(-n \le X \le x) = \int_{-n}^{x} f(t)\,dt \quad , \text{ då } -n < x$$

"kontin." för se 1.3. $\downarrow$ $\quad\downarrow n \to \infty.$

Låt $n \to \infty$

$$P(X \le x) = \int_{-\infty}^{x} f(t)\,dt$$

$\leftarrow$ (<u>orsak</u> $(-\infty, x] = \bigcup_{n\in N} (-n, x]$ växande union)

2. påståendet: Analys II $\boxtimes$

egenskaper.

<u>Sats 2.12</u> $X: \Omega \to \mathbb{R}$ s.v. med <u>kontinuerlig</u> fördelning, frekvensf. f.

Då

(i) $P(\{\omega : X(\omega) = x_0\}) = 0 \quad \forall x_0 \in \mathbb{R}$ $\quad$ $\boxed{\text{OBS: inga frekvens se!}}$

(ii) $P(a < X < b) = P(a \le X < b) = P(a < X \le b) = P(a \le X \le b)$
$= \int_a^b f(t)\,dt = F(b) - F(a)$, där $F(a) = F_X(a)$ fördeln.fkt. till X.

Inte 2015

Bevis (i) motsvarar egenskapen $\int_{x_0}^{x_0} f(t)\,dt = 0$ för integraler

idén : $n \geq 1 \Rightarrow$

$$P\left(x_0 - \tfrac{1}{n} \leq X \leq x_0\right) = \int_{x_0-\frac{1}{n}}^{x_0} f(t)\,dt$$

$n \to \infty$ $\downarrow$ $\downarrow$ f styckvis kont. (DI).

$$P(X = x_0) = 0.$$

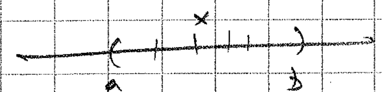
$$\{\omega : X(\omega) = x_0\} = \bigcap_{n=1}^{\infty} \underbrace{\{\omega : x_0 - \tfrac{1}{n} \leq X(\omega) \leq x_0\}}_{A_n} \Rightarrow A_{n+1} \subset A_n \subset \dots$$ (+ kontinuitet för sl).

(ii) : följer från (i) $\square$

Exempel på kontinuerliga fördelningar $(\Omega, \mathcal{F}, P)$ sl-rum.

(A) Likformig fördelning

intuition : välj ett reellt tal x "på måfå" från intervallet (a,b) så att varje tal har lika stor sl



S.v. $X : \Omega \to \mathbb{R}$ är likformigt fördelad på (a,b) om X har frekvensfunktionen

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & x \leq a \text{ eller } x \geq b. \end{cases}$$

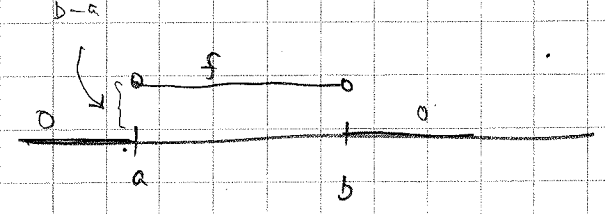
Beteckna:

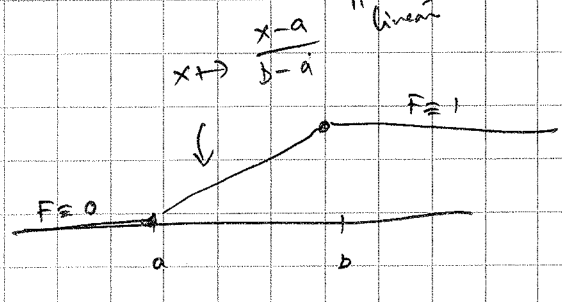
$$X \sim \text{Tas}(a,b)$$

säg även : X bestämmer likformig sl-fördelning på Ω

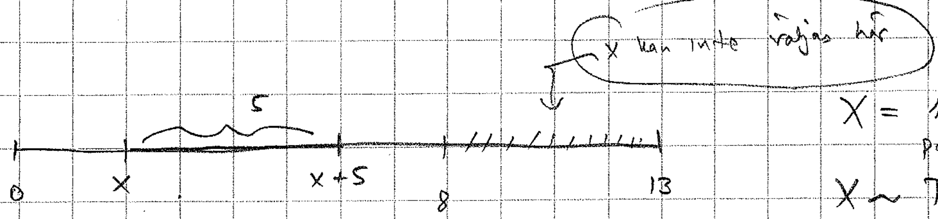
X har fördelningsfunktionen $F = F_X$:

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x < b, \\ 1, & x \geq b \end{cases}$$



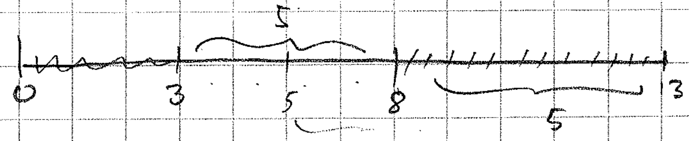


Exempel 2.13 En bil med längden 5 m. parkeras på måfå på en gata med längden 13 m. Vilken är Sl att en annan bil med längden 5 m. kan parkera på samma gata?



$X =$ 1. koordinaten x på 1. bilen $\Rightarrow$

$X \sim \text{Tas}(0,8)$.



$$\Rightarrow P(\text{2. bilen kan parkera}) = P\left(\{0 \le X \le 3\} \cup \{5 \le X \le 8\}\right)$$

$$\overset{\text{disjunkta}}{=} P(0 \le X \le 3) + P(5 \le X \le 8) = \underbrace{\int_0^3 \frac{1}{8}\,dx}_{\frac{3}{8}} + \underbrace{\int_5^8 \frac{1}{8}\,dx}_{\frac{3}{8}} = \frac{3}{4}$$

□

5/2/201

OBS samband mellan $\text{Tas}(0,1)$ och $\text{Tas}(a,b)$: $X: \Omega \to \mathbb{R}$ s.v.

$$X \sim \text{Tas}(0,1) \Rightarrow (b-a)X + a \sim \text{Tas}(a,b).$$

B. Exponentfördelningen

$\lambda > 0 \Rightarrow f(x) = \lambda e^{-\lambda x},\ x > 0$ uppfyller villkoren för frekvensf.:

$$\int_0^\infty \lambda e^{-\lambda x}\,dx = \lim_{c\to\infty} \int_0^c \lambda e^{-\lambda x}\,dx = \lim_{c\to\infty}\left(-\frac{1}{\lambda}\Big/_0^c \lambda e^{-\lambda x}\right)$$

$$= \lim_{c\to\infty}\Big[1 - \underbrace{e^{-\lambda c}}_{\frac{1}{e^{\lambda c}} \xrightarrow[c\to\infty]{} 0}\Big] = 1.$$

S.v. $X: \Omega \to \mathbb{R}$ är exponentfördelad med parameter λ (där $\lambda > 0$), om X har kontinuerlig fördeln. med frekvensfunktionen

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

$\boxed{\text{Beteckna: } X \sim \text{Exp}(\lambda)}$

Fördelningsfunktionen $F = F_X$:

$$\begin{cases} x \le 0 \Rightarrow F(x) = 0 \\ x > 0 \Rightarrow F(x) = \int_0^x \lambda e^{-\lambda t}\,dt = -\Big/_0^x \lambda e^{-\lambda t} = 1 - e^{-\lambda x} \end{cases}$$