F3 STOKASTISKA VARIABLER

Kap. 3.1–3.9

Linnéa Gustafsson
linneag2@kth.se

- Diskreta och kontinuerliga stokastiska variabler
- Exempel på fördelningar

STOKASTISKA VARIABLER

En stokastisk variabel har alltid ett tal som utfall

- En stokastisk variabel betecknas med stor bokstav, t.ex. X
- Ett utfall betecknas med liten bokstav, t.ex. x, k etc.

★ En diskret stokastisk variabel kan anta ett ändligt eller ett oändligt uppräkneligt antal värden

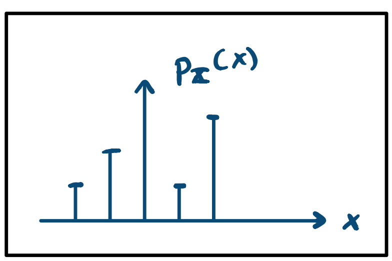
FUNKTIONER

★ Def Sannolikhetsfunktionen för $X = p_X(x)$

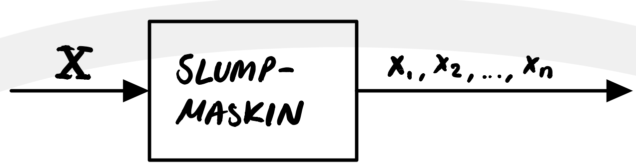
$$p_X(x) = \underbrace{P(X = x)}_{\text{Sannolikheten att } X \text{ antar värdet } x}$$

När sannolikhetsfunktionen söks så ska $p_X(x)$ anges för alla x.

$p_X(x)$ ritas upp med ett stolpdiagram:



$$\sum_{\substack{\text{alla}\\\text{utfall}\\x}} p_X(x) = 1$$



Men egentligen:

Man fixerar först x och räknar sedan ut $p_X(x)$

Ex Kasta en tärning

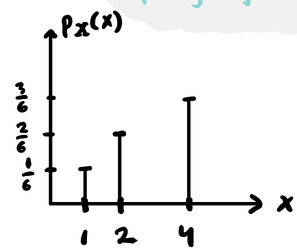
etta ⇒ vinst på 1 kr
tvåa, trea ⇒ — " — 2 kr
fyra, femma, sexa ⇒ — " — 4 kr

X = vinsten i kr

$\Omega = \{1, 2, 4\}$

$p_X(1) = \frac{1}{6}, \quad p_X(2) = \frac{2}{6}, \quad p_X(4) = \frac{3}{6}$

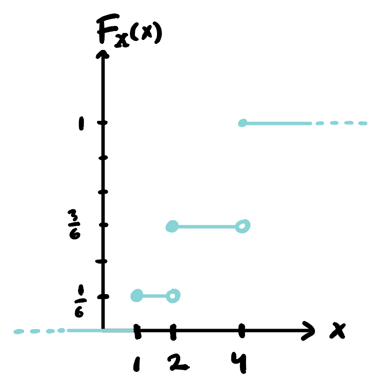
Som en vanlig funktion men här måste man ta ut varje sannolikhet för sig (ej som t.ex. $y=x^2$)



Mycket nytta av denna i det kontinuerliga fallet

Def Fördelningsfunktion för X skrivs $F_X(x) = P(X \leq x)$

I det diskreta fallet blir $F_X(x) = \sum_{\substack{\text{alla} \\ \text{utfall} \\ \leq x}} p_X(x)$



I vårt ex blir:

$F_X(1) = p_X(1) = \frac{1}{6}$

$F_X(2) = p_X(1) + p_X(2) = \frac{1}{6} + \frac{2}{6} = \frac{3}{6}$

$F_X(4) = p_X(1) + p_X(2) + p_X(4) = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} = 1$

- $0 \leq \underbrace{F_X(x) = P(X \leq x)}_{\text{sannolikhet} \Rightarrow \geq 0, \leq 1} \leq 1$
- $F_X(-\infty) = P(X \leq -\infty) = 0$
- $F_X(\infty) = P(X \leq \infty) = 1$
- $F_X(x)$ avtar aldrig

DISKRETA FÖRDELNINGAR

3 Diskreta fördelningar

Binomialfördelningen

X är Bin(n,p) om $p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, 1, ..., n$, där $0 < p < 1$.
$E(X) = np, \quad V(X) = np(1-p)$

"För-första-gången"-fördelningen

X är ffg(p) om $p_X(k) = p(1-p)^{k-1}$, $k = 1, 2, 3, ...$, där $0 < p < 1$.
$E(X) = \frac{1}{p}, \quad V(X) = \frac{1-p}{p^2}$

Hypergeometriska fördelningen

X är Hyp(N, n, p) om $p_X(k) = \dfrac{\binom{Np}{k}\binom{N(1-p)}{n-k}}{\binom{N}{n}}$, $0 \leq k \leq Np$,
$0 \leq n-k \leq N(1-p)$, där N, Np och n är positiva heltal samt $N \geq 2$, $n < N$,
$0 < p < 1$. $E(X) = np, \quad V(X) = \frac{N-n}{N-1} np(1-p)$

Poissonfördelningen

X är Po(μ), där $\mu > 0$, om $p_X(k) = \frac{\mu^k}{k!} e^{-\mu}$, $k = 0, 1, 2, ...$
$E(X) = \mu, \quad V(X) = \mu$

★ Tvåpunktsfördelningen

- Speciellt: **Bernoullifördelningen**

$\Omega = \{0, 1\}$

$p_X(1) = p, \quad p_X(0) = 1 - p$

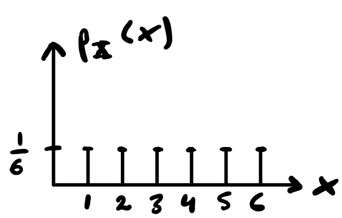
Användning:

- Varje gång försöket lyckas (p) sätter vi 1
- — " — misslyckas $(1-p)$ — " — 0

Den vanligaste är den kontinuerliga

★ Den diskreta likformiga fördelningen

Här är varje möjligt utfall lika sannolikt



"För-första-gången"-fördelningen (se § 3 i F.S.)

På engelska: Geometrical distribution

Här är X antal försök t.o.m. det första lyckade där vi antar att vi i varje försök har sannolikheten p att lyckas.

A_i – vi lyckas i försök nr i

$P(A_i) = p$

$P(X=1) = P(A_1) = p$

$P(X=2) = P(A_1^* \cap A_2) = [\text{ober.}] = P(A_1^*) \cdot P(A_2) = (1-p) \cdot p$

$P(X=3) = P(A_1^* \cap A_2^* \cap A_3) = [\dots] = (1-p) \cdot (1-p) \cdot p$

$p_X(k) = (1-p)^{k-1} \cdot p$

$X = 1, 2, 3, \dots$

Geometrisk fördelning

Här är X antal misslyckade försök

$X = 0, 1, 2, \dots$

$p_X(0) = p, \quad p_X(1) = (1-p) \cdot p$ "Man backar ett steg" [jmf "f-f-g"-fördeln.]

$p_X(k) = (1-p)^k \cdot p$

Binomialfördelningen

På ett sätt samma situation som för "f-f-g"-fördeln. men slutar inte 1:a ggn vi lyckas, utan efter n försök

Vi gör n st försök.

Varje gång har vi samma sannolikhet p att lyckas.

X = antal gånger vi lyckas

$X \in \text{Bin}(n, p)$ X tillhör binomialfördelningen (n, p)

Ex Vi gör 7 försök

Vad är $P(X=3)$ om $X \in Bin(7, p)$?

misslyckas   lyckas

En variant: $\circ \bullet \bullet \circ \circ \bullet \circ$

$(1-p)\ p\ p\ (1-p)\ (1-p)\ p\ (1-p) = p^3(1-p)^4$

En annan: $\bullet \circ \circ \circ \bullet \circ \bullet$

$p\ (1-p)\ (1-p)\ (1-p)\ p\ (1-p)\ p = p^3(1-p)^4$

Antal sätt att plocka in tre lyckade på 7 platser $= \binom{7}{3}$

$\Rightarrow P(X=3) = \binom{7}{3} p^3 (1-p)^4$

Allm när $X \in Bin(n, p)$

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

★ Hypergeometriska fördelningen

Vi drar utan återläggning.

Vi har från början N st enheter, där andelen med egenskap A är p.

Vi drar sedan n st enheter utan återläggning.

X = antal enheter med egenskap vi får.

Ex (se sid. 24 i Blom)

Vi hade v vita och s svarta kulor och vi drog n st kulor utan återläggning.

X = antal vita vi får

$$p_X(k) = \frac{\binom{v}{k}\binom{s}{n-k}}{\binom{v+s}{n}}$$

I F.S. § 3 står $X \in Hyp(N, n, p)$

$$\Rightarrow p_X(k) = \frac{\binom{Np}{k}\binom{N(1-p)}{n-k}}{\binom{N}{n}} \quad *$$

I kulfallet gäller ju att

$$N = v + s$$

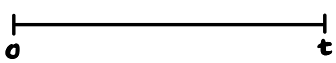
$$p = \frac{v}{v+s}$$ andelen vita vi har från början

$$p_X(k) = \frac{\binom{v}{k}\binom{s}{n-k}}{\binom{v+s}{n}} = \frac{\binom{\overbrace{(v+s)}^{n}\overbrace{\frac{v}{v+s}}^{p}}{k}\binom{v+s-v}{n-k}}{\binom{N}{n}} = \frac{\binom{Np}{k}\binom{\overbrace{v+s}^{N}(1-\overbrace{\frac{v}{v+s}}^{p})}{n-k}}{\binom{N}{n}} =$$

$$* = \frac{\binom{Np}{k}\binom{N(1-p)}{n-k}}{\binom{N}{n}}$$

Poissonfördelningen

Anta att vi har ett intervall $(0, t]$



Vi antar att det i varje tidpunkt mellan 0 och t är lika stor sannolikhet att en händelse inträffar, oberoende av om det redan inträffat en händelse eller ej.

X = antal händelser som inträffat på intervallet $[0, t]$.

Pg.a. fixerat tiden

Vi antar att händelserna inträffar med intensiteten λ.

$$\lambda = \frac{\text{händelser}}{\text{tidsenhet}}$$

i F.S. §3 gäller att $\lambda t = \mu \Rightarrow p_X(k) = \frac{\mu^k}{k!}e^{-\mu}$
(μ antal händelser i genomsnitt under tidsint.)

$$\Rightarrow \underbrace{p_X(k)}_{\text{antalet händelser i tidsintervallet } = k} = \frac{(\lambda t)^k}{k!}e^{-\lambda t} \qquad k = 0, 1, 2, \ldots$$