FORELESNING 5

EIVIND ERIKSEN

JAN 31 2012

FLE 3719

MATEMATIKK VALGFAG

PLAN:

1. Oppgaveark 2 (Oppg. 2.30, 2.31, 2.32 fra [R] utsettes til neste uke)
2. Poisson fordeling [R] 2.2
3. Kontinuerlige fordelinger [R] 2.3
   - Generelt, Uniform fordeling.

---

1. Ingen oppgaver er vanskelige?
2. Poisson fordelingen.

Defn: En diskret stokastisk variabel X med mulige verdier $x = 0, 1, 2, 3, \ldots$ er Poisson med parameter λ hvis

$$f_X(i) = P(X = i) = \frac{\lambda^i}{i!} e^{-\lambda}, \quad i = 0, 1, 2, \ldots$$

Typisk bruk:

X = antall forekomster av en hendelse i et bestemt tidsintervall

λ = forventet antall forekomster

Poisson-fordeling viktig pga

(1) Mange anvendelser i seg selv

(2) Tilnærmer binomisk fordeling godt når n er stor og p er liten:

$$\underbrace{\binom{n}{i} \cdot p^i \cdot (1-p)^{n-i}}_{\substack{p(X=i) \\ \text{når } X \text{ er binomisk} \\ \text{med parametre } (n,p)}} \simeq \underbrace{\frac{\lambda^i}{i!} e^{-\lambda}}_{\substack{P(X=i) \\ \text{når } X \text{ er Poisson} \\ \text{med parameter } \lambda}}$$

$\boxed{\text{tilnærmet like når } n \text{ er stor, } p \text{ liten og } \lambda = n \cdot p}$

Eks: X binomisk med $n=100$, $p=0.01$

$$p(X=2) = \binom{100}{2} \cdot 0.01^2 \cdot 0.99^{98} \simeq \underline{0.18486}$$

$\Downarrow$

Y Poisson, $\lambda = np = 100 \cdot 0.01 = 1$

$$p(Y=2) = \frac{\lambda^2}{2!} e^{-\lambda} = \frac{1^2}{2!} e^{-1} = \frac{1}{2e} \simeq \underline{0.18394}$$

[Previous] [Next] [Up] [Top] **Categorical Data Analysis with Graphics**
Michael Friendly

Part 1: Plots for discrete distributions

Contents

Discrete frequency distributions often involve counts of occurrences such as accidental fatalities, words in passages of text, or blood cells with some characteristic. Typically such data consist of a table which records that *n sub k* of the observations pertain to the basic outcome value *k* , *k = 0 , 1, ...* .

The table below shows two such data sets:

- von Bortkiewicz's (1898) data on death of soldiers in the Prussian army from kicks by horses and mules. The data pertain to 10 army corps, each observed over 20 years. In 109 corps-years, no deaths occurred; 65 corps-years had one death, etc. (Figure 1)
- Mosteller & Wallace's (1964) data on the occurrence of the word *may* in 262 blocks of text (each about 200 words long) from issues of the *Federalist Papers* known to be written by James Madison. In 156 blocks, the word *may* did not occur; it occurred once in 63 blocks, etc. (Figure 2)

```
Deaths by Horsekick              Occurrences of 'may'

   k     nk                          k     nk
  ----------                        ----------
   0     109                         0     156
   1      65                         1      63
   2      22                         2      29
   3       3                         3       8
   4       1                         4       4
        -----                        5       1
        N=200                        6       1
                                          -----
                                          N=256
```

$$P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$P(X=0) = \frac{\lambda^0}{0!} e^{-\lambda} = e^{-\lambda}$$

$$P(X=1) = \frac{\lambda^1}{1!} e^{-\lambda} = \lambda e^{-\lambda}$$

$$P(X=2) = \frac{\lambda^2}{2!} e^{-\lambda} = \frac{\lambda^2}{2} \cdot e^{-\lambda}$$

$\lambda = 0.6$

$P(X=0) \cdot 200 \sim 109.8$

$P(X=1) \cdot 200 \sim 65.9$

$P(X=2) \cdot 200 \sim 19.8$

$P(X=3) \cdot 200 \sim 4.0$

$P(X=4) \cdot 200 \sim 0.6$

$$\lambda \simeq \frac{65/200}{109/200} = \frac{65}{109} \simeq \underline{0.6}$$

since $\frac{P(X=1)}{P(X=0)} = \frac{\lambda e^{-\lambda}}{e^{-\lambda}} = \lambda$

Er Poisson-fordelingen en sannsynlighetsfordeling?

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

i) Er $\frac{\lambda^i}{i!} e^{-\lambda} \geq 0$? Ja, for $\lambda > 0$.

ii) Er $\sum_{i=0}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda} = 1$?

$$\sum_{i=0}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda} = \frac{\lambda^0}{0!} \underline{e^{-\lambda}} + \frac{\lambda^1}{1!} \underline{e^{-\lambda}} + \frac{\lambda^2}{2!} \underline{e^{-\lambda}} + \ldots$$

$$= e^{-\lambda} \cdot \left(1 + \lambda + \frac{\lambda^2}{2} + \frac{\lambda^3}{3!} + \ldots\right)$$

$$= e^{-\lambda} \cdot e^{\lambda} = 1$$

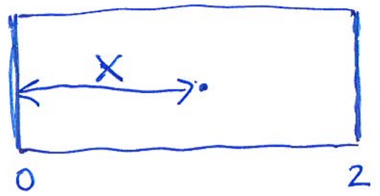
<u>Bruker at</u>: $e^{\lambda} = 1 + \lambda + \frac{\lambda^2}{2} + \frac{\lambda^3}{3!} + \ldots$

(kjent formel for eksponensialfunksjonen)

3 Kontinuerlige fordelinger [R] 2.3

X er <u>kontinuerlig</u> stokastisk variabel hvis mengden av mulige verdier for X er kontinuerlig (dvs. den inneholder et intervall).

<u>Eks</u>:



Vi kaster en dartpil

X er antall meter fra venstre vegg til det punktet vi treffer.

Mulige verdier for X : $[0,2]$

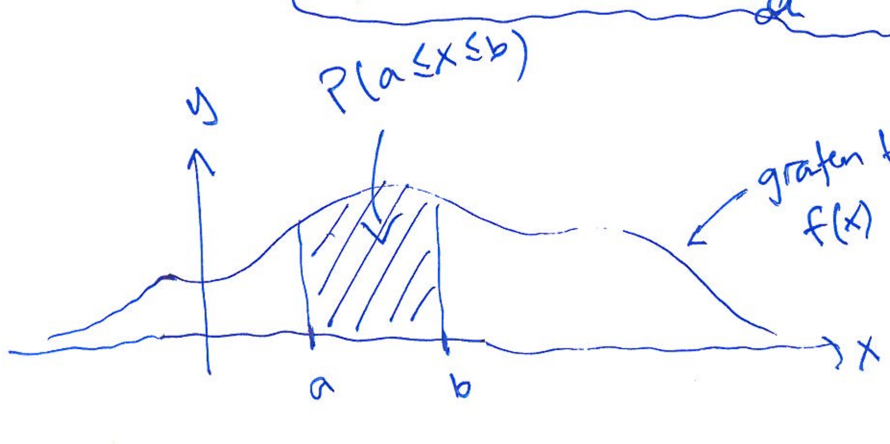
Hva $p(X=1)$? $p(X=1)=0$!

$p(X\leq 1)$? $p(X\leq 1)=1/2$ hvis fordelingen er <u>uniform</u>. $\frac{1-0}{2-0}=\frac{1}{2}$

$p(0.4\leq X\leq 0.5)$? $p(0.4\leq X\leq 0.5)=\frac{0.5-0.4}{2}=\underline{0.05}$

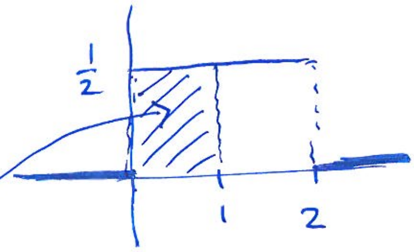
<u>Defn</u>: X er en kontinuerlig fordelt stokastisk variabel hvis det finnes en sannsynlighetstetthet $f_X(x)=f(x)$ slik at

$$\boxed{p(a\leq X\leq b)=\int_a^b f(x)\,dx}$$



Eks: Uniform stokastisk variabel

$$f(x) = \begin{cases} \frac{1}{2}, & 0 \le x \le 2 \\ 0, & \text{ellers} \end{cases}$$



Sannsynlighetstetthet for pilkast-eksempelet

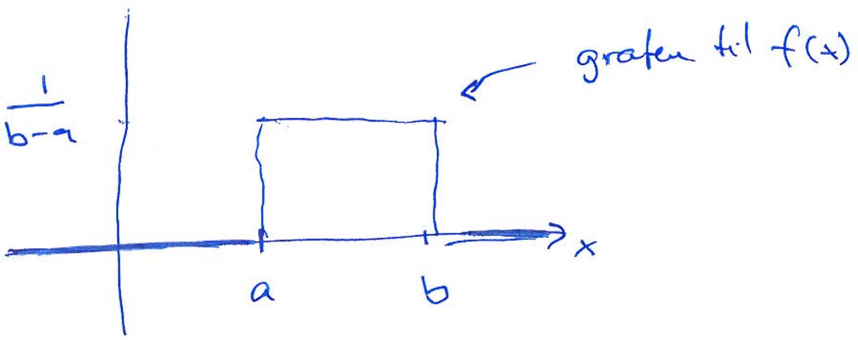
$$P(X \le 1) = P(0 \le X \le 1) = \int_0^1 f(x)\,dx = \int_0^1 \frac{1}{2}\,dx = 1 \cdot \frac{1}{2} = \underline{\underline{\frac{1}{2}}}$$

alt.

$$= \left[\frac{1}{2}x\right]_0^1 = \frac{1}{2}\cdot 1 - \frac{1}{2}\cdot 0 = \underline{\underline{\frac{1}{2}}}$$

Defn. X er uniform på intervallet $[a,b]$ hvis

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b] \\ 0, & \text{ellers} \end{cases}$$



Oppsummering: Diskret vs Kontinuerlig

	Diskret	Kontinuerlig
Sannsynlighets-tetthet:	$f(x) = P(X=x)$	$f(x)$ $P(X=x) = 0$ for alle x
Sannsynlighet:	$P(a \leq X \leq b) = \sum_{a \leq x \leq b} P(X=x)$	$P(a \leq X \leq b) = \int_a^b f(x)\,dx$
Kumulativ fordelingsfunksjon:	$F(b) = P(X \leq b) = \sum_{i \leq b} P(X=i) = \sum_{x \leq b} f(x)$	$F(b) = P(X \leq b) = \int_{-\infty}^{b} f(x)\,dx$
		$f(x) = F'(x)$
Krav til sannsynlighets-tetthet $f(x)$	i) $f(x) \geq 0$ for alle x ii) $\sum_{x_i} f(x_i) = 1$	i) $f(x) \geq 0$ for alle x ii) $\int_{-\infty}^{\infty} f(x)\,dx = 1$