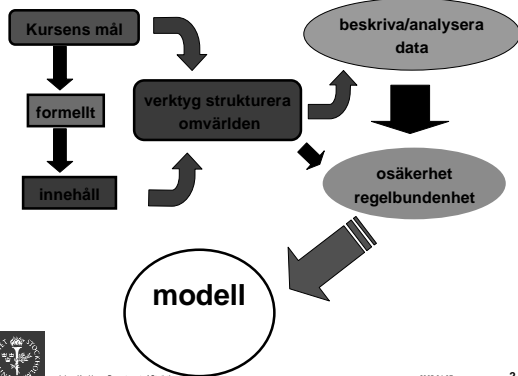




Johan Koskinen, Statistiska institutionen, Stockholms universitet

Finansiell statistik, vt-05

F5 Diskreta variabler



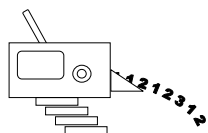
Johan Koskinen, Department of Statistics

2005-04-07

2

Modeller

- Beskriver sannolikheter för att observera saker vi kan observera
- Bernoulli
- Binomial
- Hypergeometrisk fördelning
- Poisson



Johan Koskinen, Department of Statistics

2005-04-07

3

Fördelningar: Hypergeometrisk

I en skål finns 2 gröna marker och 3 röda. Drag två marker. Vad är sannolikheten att vi fått exakt 1 grön marker?

Låt X vara antal gröna marker vid två dragningar (u.å.)



Variabeln X antar värdena: 0, 1, 2.

Variabeln X är hypergeometriskt fördelad och har slh-funktion av formen

$$P(A) = \frac{g(A)}{m}$$

hur beräknar vi antalet gynnsamma utfall och totala ant utfall?



Johann Koskinen, Department of Statistics

2005-04-07

4

Fördelningar: Hypergeometrisk

Det totala antalet sätt vi kan dra två marker på är

$$\binom{5}{2} = \frac{5!}{(5-2)!2!} = \frac{(5)(4)}{(2)(1)} = 10$$



Antal utfall som är gynnsamma för $x = 0$:

antal sätt röd i 1:a röd i 2:a;

antal sätt välja 2 av tre

$$\binom{3}{2} = \frac{3!}{1!2!} = 3$$

$$p(0) = P(X=0) = \frac{\binom{3}{2}}{\binom{5}{2}} = \frac{3}{10}$$



Johann Koskinen, Department of Statistics

2005-04-07

5

Fördelningar: Hypergeometrisk

Antal utfall som är gynnsamma för $x = 1$:

antal sätt grön i 1:a och röd i 2:a +

antal sätt grön i 2:a och röd i 1:a

Notera: ordning ej av betydelse, alltså



Antal utfall som är gynnsamma för $x = 1$:

antal sätt dra 1 grön och 1 röd

antal sätt dra 1 grön \times antal sätt dra 1 röd

$$\binom{2}{1}\binom{3}{1} = \frac{2!}{1!1!} \frac{3!}{1!2!} = 6$$

$$p(1) = P(X=1) = \frac{\binom{2}{1}\binom{3}{1}}{\binom{5}{2}} = \frac{6}{10}$$



(alternativ: se dragningarna som ober. försök...)

Johann Koskinen, Department of Statistics

2005-04-07

6

Fördelningar: Hypergeometrisk

Antal utfall som är gynnsamma för $x = 2$:
antal sätt grön i 1:a och grön i 2:a

$$\binom{2}{2} = \frac{2!}{0!2!} = 1$$



$$p(2) = P(X = 2) = \frac{\binom{2}{2} \binom{3}{2}}{\binom{5}{2}} = \frac{1}{10}$$



Johann Koskinen, Department of Statistics

2005-04-07

7

Fördelningar: Hypergeometrisk

Om X är Hypergeometriskt fördelat med parametrarna N , n och h skriver vi
 $X \in \text{Hyp}(N, h, n)$

Den Hypergeometriskt fördelad X antar värdena:
 $h - (N - n)$, $h - (N - n) + 1$, ..., $\min(h, n)$.
Har sannolikhetsfunktionen:

$$p(x) = \frac{\binom{h}{x} \binom{N-h}{n-x}}{\binom{N}{n}}$$

Tolkning:
sannolikheten att vid dragning av n objekt utan återläggning dra x st element det finns h av totalt N



Johann Koskinen, Department of Statistics

jfr binomial: med återläggning $p = h/N$

Fördelningar: Poisson

Siméon Denis Poisson (1781-1840):
antal Preussiska soldater ihjälsparkade av mulor

Typiskt: poissonfördelad variabel X beskriver antal

$x = 0, 1, 2, 3, 4, 5, 6, \dots$



- Antal dgr under ett år då en 50-punkts förändring sker på Dow Jones
- Antal samtal till en tfn-växel under en given period
- Antal mikrober på ett glas



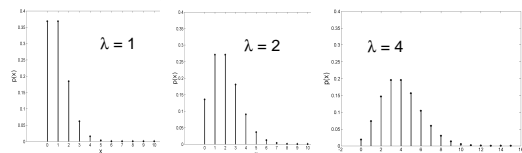
Johann Koskinen, Department of Statistics

2005-04-07

9

Fördelningar: Poisson

Slh-funktionen för en poissonfördelad variabel X
beror av en parameter $\lambda > 0$



Johannes Koskinen, Department of Statistics

2005-04-07

10

Fördelningar: Poisson

En poissonfördelad variabel X
 $X \in \text{Poisson}(\lambda)$
har sannolikhetsfunktion

$$p(x) = e^{-\lambda} \frac{\lambda^x}{x!}$$

Antag att för antal samtal X till en tfn-växel under en 10 min
 $X \in \text{Poisson}(4)$
vad är sannolikheten att det kommer in exakt 2 samtal under tio-minuter.?



Johannes Koskinen, Department of Statistics

2005-04-07

11

Fördelningar: Poisson

$X \in \text{Poisson}(4)$
slh att det kommer in exakt 2 samtal under tio-minuter.

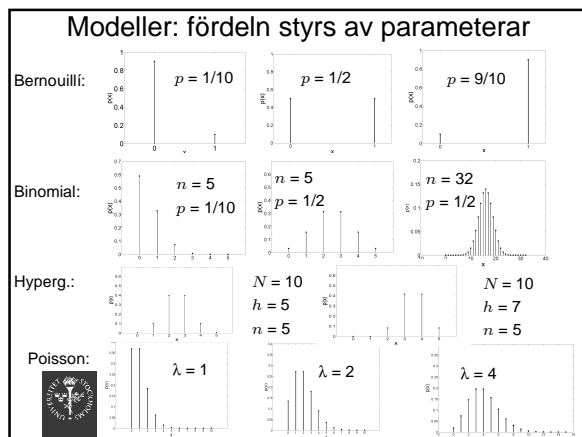
$$\begin{aligned} P(X=2) &= p(2) \\ &= e^{-4} \frac{4^2}{2!} \\ &= (0,0183) \left(\frac{16}{2} \right) = 0,1465 \end{aligned}$$



Johannes Koskinen, Department of Statistics

2005-04-07

12



Fördelningsfunktionen

Tag exemplet $X \in \text{Binomial}(32, 1/2)$
fördelningen ger oss:

$P(X=10)$

$$p(10) = \binom{32}{10} (0.5)^{10} (0.5)^{32-10}$$


$$= 64512240 \times 2^{-32} = 0,0150$$

kanske mer intressant:

$$P(X \leq 10) = P(\{X=0\} \cup \{X=1\} \cup \dots \cup \{X=10\})$$

$$= P(X=0) + P(X=1) + \dots + P(X=10)$$

$$= p(0) + p(1) + \dots + p(10)$$

 Johan Koskinen, Department of Statistics 2005-04-07 14

Fördelningsfunktionen


Om X är en diskret stokastisk variabel som antar värdena x_1, x_2, \dots, x_n , där $x_1 < x_2 < \dots < x_n$ kallar vi

$$F(x) = P(X \leq x)$$

för den stokastiska variabelns fördelningsfunktion.

Fördelningsfunktionen kan alltså skrivas

$$F(x_k) = \sum_{\text{alla } x \leq x_k} P(X=x) = \sum_{i=1}^k p(x_i)$$

 Johan Koskinen, Department of Statistics 2005-04-07 15

Fördelningsfunktionen

Antag att för antal samtal X till en tfn-växel under en 10 min
 $X \in \text{Poisson}(4)$
 vad är slh att det kommer in högst 2 samtal under tio-minper.?

$$P(X \leq 2)$$

$$\begin{aligned} F(2) &= p(0) + p(1) + p(2) \\ &= e^{-4} \frac{4^0}{0!} + e^{-4} \frac{4^1}{1!} + e^{-4} \frac{4^2}{2!} \\ &= (0,0183)(1 + 4 + 8) = 0,2381 \end{aligned}$$



Johannes Koskinen, Department of Statistics

2005-04-07

16

Fördelningsfunktionen

$$X \in \text{Bernoulli}(p) \quad F(x) = \sum_{j=0}^x p(j) = \sum_{j=0}^x p^j q^{1-j}$$

$$X \in \text{Binomial}(n, p) \quad F(x) = \sum_{j=0}^x p(j) = \sum_{j=0}^x \binom{n}{j} p^j q^{n-j}$$

$$X \in \text{Hyp}(N, h, n) \quad F(x) = \sum_{j=0}^x p(j) = \binom{N}{n}^{-1} \sum_{j=0}^x \binom{h}{j} \binom{N-h}{n-j}$$

$$X \in \text{Poisson}(4) \quad F(x) = \sum_{j=0}^x p(j) = \sum_{j=0}^x e^{-4} \frac{4^j}{j!}$$



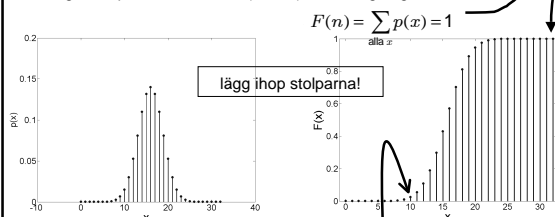
Johannes Koskinen, Department of Statistics

2005-04-07

17

Fördelningsfunktionen

Tag exemplet $X \in \text{Binomial}(32, 1/2)$ fördelningen ger oss:



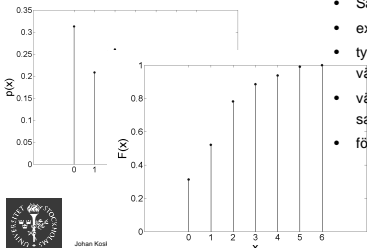
Johannes Koskinen, Department of Statistics

2005-04-07


18

Väntevärde

Om X är en diskret stokastisk variabel
hur beskriva dess sannolikhetsfördelning?



- Sammanfatta
- ex: mest sannolikt värde
- tyngdpunkt/mest centralt värde
- väga värden med sannolikheterna
- förväntad "vinst"



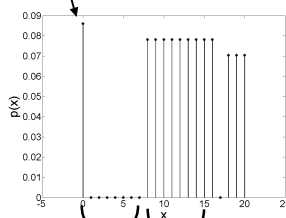
Johannes Koski

2005-04-07

19


Väntevärde

Mest sannolika enskilda värde på X är $x = 0$



sannolikheten att hamna här <

sannolikheten att hamna här



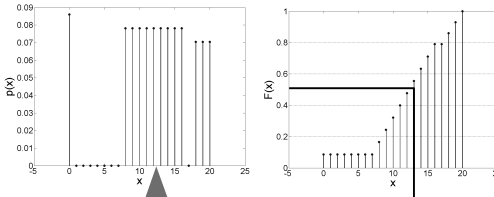
Johannes Koski

2005-04-07

20

Väntevärde


tyngdpunkt/mest centralt värde



12,45

"minsta avståndet" till alla värden

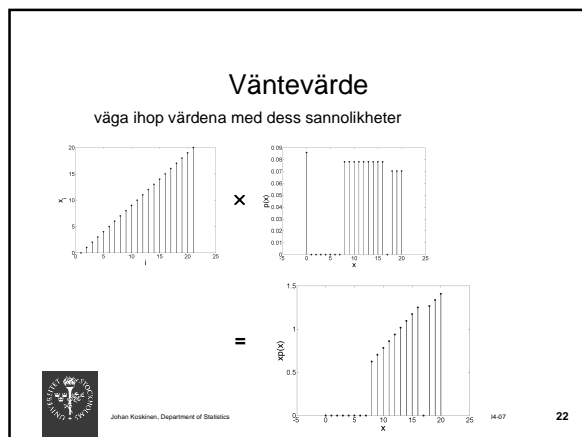
$P(X \leq 13) \approx 1/2$



Johannes Koski, Department of Statistics

2005-04-07

21



Väntevärde

Antag att vi vinner $x = 1$ dollar med slh p
vi vinner $x = 0$ dollar med slh $1 - p$

Hur stor vinst kan vi förvänta oss?

Svar: $0 \times (1 - p) + 1 \times p = p$

Om vi har frekventisttolkning av sannolikheter:
 p den genomsnittliga vinsten vi skulle få om vi spelade
"många" ggr

Johannes Koskinen, Department of Statistics

2005-04-07

23

Väntevärde: definition

Om X är en diskret stokastisk variabel som antar värdena
 x_1, x_2, \dots, x_n
med sannolikhetsfunktionen
 $p(x_i)$
är väntevärdet för X :

$$E(X) = \sum_{\text{alla } x} xP(X = x) = \sum_{i=1}^n x_i p(x_i)$$

Johannes Koskinen, Department of Statistics


2005-04-07

24

Väntevärde för funktioner av s.v.

Om X är en diskret stokastisk variabel som antar värdena x_1, x_2, \dots, x_n med sannolikhetsfunktionen $p(x_i)$ och f är en funktion av X är väntevärdet för $f(X)$:

$$E(f(X)) = \sum_{\text{alla } x} f(x)P(X = x) = \sum_{i=1}^n f(x_i)p(x_i)$$



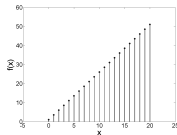
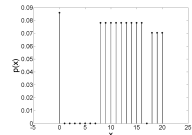
Johannes Koskinen, Department of Statistics

2005-04-07

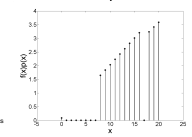
25

Väntevärde för funktioner av s.v.


Om X diskret stokastisk variabel
 $f(x) = 1 + 2,5x$ är en funktion av X ; är väntevärdet för $f(X)$:

×




=



summan av staplarna:

32



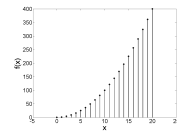
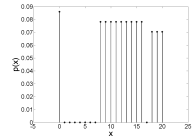
Johannes Koskinen, Department of Statistics

2005-04-07

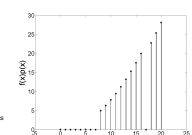
26

Väntevärde för funktioner av s.v.


Om X diskret stokastisk variabel
 $f(x) = x^2$ är en funktion av X ; är väntevärdet för $f(X) = E(X^2)$:

×




=



summan av staplarna:

182



Johannes Koskinen, Department of Statistics

2005-04-07

27

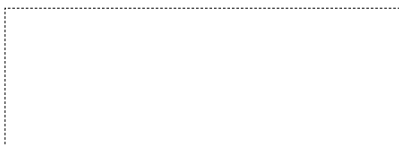
Väntevärde för funktioner av s.v.

Väntevärdet är en lineär operator:

Om X är en diskret stokastisk variabel och

$$Y = a + bX$$

är $E(Y) = a + bE(X)$



Johannes Koskinen, Department of Statistics

2005-04-07

28

Varians

Om X är en diskret stokastisk variabel som antar värdena

x_1, x_2, \dots, x_n

med sannolikhetsfunktionen

$p(x_i)$

är variansen för X :

$$\begin{aligned} E([X - E(X)]^2) &= \sum_{\text{alla } x} (x - E(X))^2 P(X=x) \\ &= \sum_{i=1}^n (x_i - E(X))^2 p(x_i) \end{aligned}$$



Johannes Koskinen, Department of Statistics

2005-04-07

29
