

# Föreläsning 4: Kontinuerliga fördelningar

## Matematisk statistik

David Bolin  
Chalmers University of Technology  
April 2, 2015



# Slumpvariabler

Ofta vill vi beskriva utfallet av ett slumpmässigt försök som ett numeriskt värde och man kan då beskriva försöket med en slumpvariabel

## Slumpvariabel

En stokastisk variabel (slumpvariabel)  $X$  är en funktion som tar element från  $\Omega$  som argument och som returnerar ett reellt tal.

Vi kommer beteckna slumpvariabler med stora bokstäver, ofta  $X$ ,  $Y$  och  $Z$ .

## Diskreta slumpvariabler

En diskret slumpvariabel kan endast anta ett ändligt eller uppräkneligt antal värden, i allmänhet någon delmängd av heltalen.

# Sannolikhetsfunktionen

## Sannolikhetsfunktion

Till en diskret stokastisk variabel  $X$  definierar vi sannolikhetsfunktionen  $p(k)$  genom  $p(k) = P(X = k)$ .

Alla funktioner är inte sannolikhetsfunktioner. Eftersom de beskriver sannolikheter måste vi ha

- $p(k) \geq 0$  för alla  $k$ .
- $\sum_{k=-\infty}^{\infty} p(k) = 1$ .

Dessa två villkor är både nödvändiga och tillräckliga för att  $p(k)$  ska vara en sannolikhetsfunktion. Vi har också att

$$P(m \leq X \leq n) = \sum_{k=m}^n p(k)$$

om  $m$  och  $n$  är heltal.

Notera att sannolikhetsfunktionen betecknas med  $f(s)$  i boken.

# Fördelningsfunktioner

## Fördelningsfunktioner

Låt  $X$  vara en diskret slumpvariabel. Dess fördelningsfunktion ges då av

$$F(x) = P(X \leq x) = \sum_{k \leq x} p_X(k),$$

För  $F(x)$  gäller att

- $F(x)$  är växande.
- $F(x) \rightarrow 1$  då  $x \rightarrow \infty$ .
- $F(x) \rightarrow 0$  då  $x \rightarrow -\infty$ .

Vi har också att:

- $P(a < X \leq b) = F(b) - F(a)$ .
- $P(X > a) = 1 - F(a)$ .
- För diskreta slumpvariabler gäller  $p(m) = F(m) - F(m - 1)$ .

# Bernoullifördelningen

Antag att vi gör ett försök som har sannolikhet  $p$  att lyckas och sätter

$$X = \begin{cases} 1, & \text{om försöket lyckas} \\ 0, & \text{om försöket misslyckas.} \end{cases}$$

Vi har då  $p(1) = p$  och  $p(0) = 1 - p$ .

Lite mer kompakt kan vi skriva fördelningen som  $p(k) = p^k(1 - p)^{1-k}$ .

## Bernoullifördelningen

En stokastisk variabel  $X$  sägs vara Bernoullifördelad om den har sannolikhetsfunktion  $p(k) = p^k(1 - p)^{1-k}$ ,  $k=0,1$ . Vi beteckar den som  $\text{Be}(p)$ .

# Binomialfördelningen

Binomialfördelningen, förkortad  $\text{Bin}(n, p)$ , är en fördelning som beskriver antalet lyckade försök av totalt  $n$  stycken Bernoulli-försök.

## Binomialfördelningen

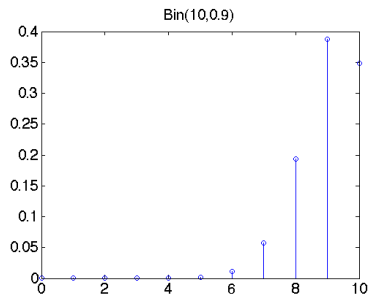
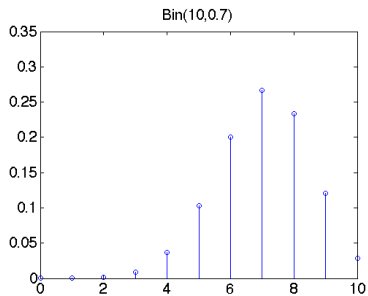
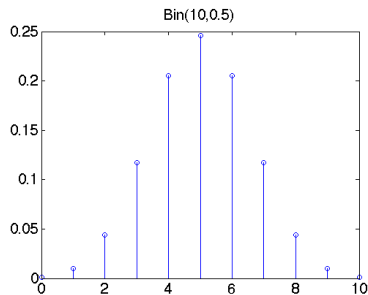
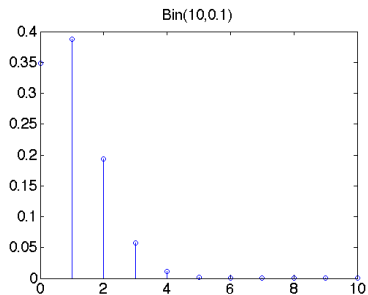
En stokastisk variabel  $X$  sägs vara binomialfördelad,  $\text{Bin}(n, p)$ , om den har sannolikhetsfunktion

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

där

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (2)$$

och  $k! = k \cdot (k-1) \cdots 2 \cdot 1$ .



# Binomialfördelningen

$\text{Bin}(n, p)$ -fördelningen kan ses som summan av  $n$   $\text{Be}(p)$  variabler.

Därför är  $\text{Bin}(1, p)$ -fördelningen är den samma som  $\text{Be}(p)$ -fördelningen

Lite mer generellt kan vi skriva upp följande sats:

## Sats

Om  $X_1$  är  $\text{Bin}(n_1, p)$  och  $X_2$  är  $\text{Bin}(n_2, p)$  samt att  $X_1$  och  $X_2$  är oberoende gäller att  $X_1 + X_2$  är  $\text{Bin}(n_1 + n_2, p)$ -fördelad.



# Den geometriska fördelningen

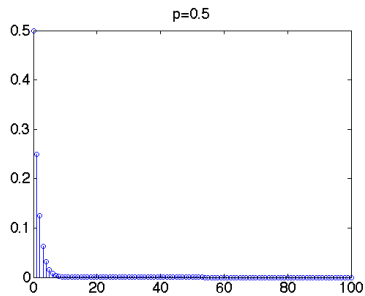
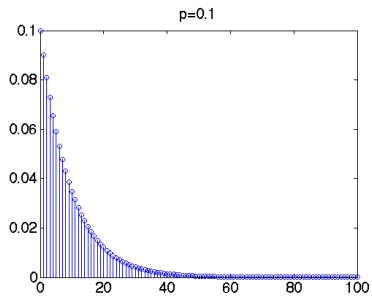
Den geometriska fördelningen beskriver antalet försök fram till det första lyckade i en serie av Bernoulli-försök.

## Geometrisk fördelningen

Slumpvariabeln  $X$  är geometriskt fördelad med parameter  $p$  om den har sannolikhetsfunktion

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

Att fördelningen ser ut som den gör inses enkelt eftersom vi för att få  $X = k$  först måste få  $k - 1$  misslyckade försök, som har sannolikhet  $1 - p$ , och sedan ett lyckat försök.



# Poissonfördelningen

Poissonfördelningen är döpt efter Simeon Denis Poisson (1781-1840). Den används ofta för att beskriva fördelningen av antalet händelser av något slag som inträffar under ett tidsintervall, på en yta, eller i en volym.

Några exempel där denna fördelning passar bra är

- Räkna antalet radioaktiva partiklar som emitteras under en minut från ett radioaktivt material.
- Registrera antalet inkommande telefonsamtal till en telefonväxel under en timme.

# Poissonfördelningen

## Poissonfördelningen

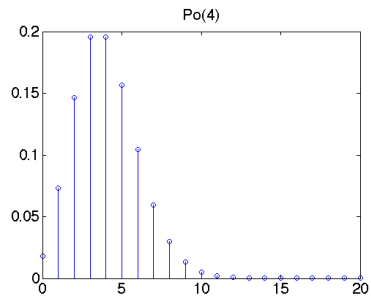
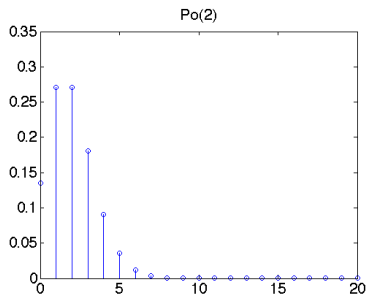
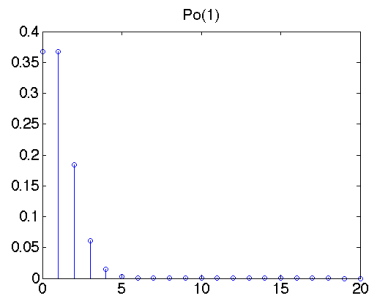
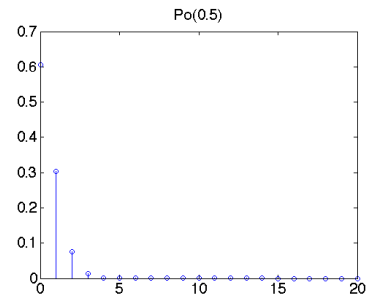
En slumpvariabel  $X$  sägs vara Poissonfördelad med parameter  $\mu$  om den har sannolikhetsfunktion

$$p(k) = \frac{\mu^k e^{-\mu}}{k!}$$

Vi använder ofta beteckningen  $\text{Po}(\mu)$  för denna fördelning.

## Summering av Poissonvariabler

Vi har vi att om  $X_1$  är  $\text{Po}(\mu_1)$  och  $X_2$  är  $\text{Po}(\mu_2)$  så gäller att  $X_1 + X_2$  är  $\text{Po}(\mu_1 + \mu_2)$ .



# Poissonfördelningen som gränsfördelning

Konceptuellt kan fördelningen ses som en gränsfördelning till binomialfördelningen om vi låter  $n$  gå mot oändligheten och  $p$  gå mot noll. Mer specifikt har vi

## Sats

Då  $n \rightarrow \infty$ ,  $p \rightarrow 0$ , och  $np \rightarrow \mu$  så gäller för ett fixt heltal  $k \geq 0$  att

$$\binom{n}{k} p^k (1-p)^{n-k} \rightarrow \frac{\mu^k e^{-\mu}}{k!} \quad (3)$$

I exemplen ovan kan vi tolka den här satsen som

- Det finns ett stort antal atomer ( $n$ ) i materialet och sannolikheten ( $p$ ) att en specifik atom ska emittera en partikel under just den tidsperioden är mycket liten.
- Ett stort antal personer kan, oberoende av varandra, ringa växeln, men för en specifik person är sannolikheten att denne ska ringa växeln mycket liten.