

Uppgift 3) Bevis för linear regression line..

Låt S representera ”Sum of errors”. Genom hela beviset är nedregränsen för summaformeln i = 1 och övre är n.

Vi vill nu minimera SSE mha partiell derivata.   
  
  
  
Sätt derivatan till lika med 0.

Substituerar (\*\*) in i (\*)

En jävla massa algebra..

Uppgift 2)

Ge definitionen för en genererande funktion?

Den genererande funktionen för den oändliga

Sekvensen (g0,g1,g2,g3,…) är:

G(x) = g0 + g1 x + g2 x^2 + g3 x^3 + …  
Summan av den oändliga geometriska serien är:

1 + x + x^2 + x^3 + … = 1 / (1-x)

Vi vill hitta sekvensen (1, 4, 9, 16, …)

(1,1,1,1, …) 🡨🡪 1 + x + x^2 + x^3 + … 🡨🡪 1 / (1-x)

d/dx(1 + x + x^2 + x^3 + …) = d/dx(1 / (1 – x))

1 + 2x + 3x^2 + 4x^3 + … = 1 / (1-x)^2

(1, 2, 3, 4, …) 🡨🡪 1 / (1-x)^2

(0, 1, 2, 3, 4, …) 🡨🡪 x / (1-x)^2

d/dx( x / (1-x)^2 ) = (1+x) / (1-x)^3

(1, 4, 9, 16, …) 🡨🡪 (1+x) / (1-x)^3

Regler för genererande funktioner…

F(x) = (f0, f1, f2, f3, …)

G(x) = (g0, g1, g2, g3, …)

Multiplikation med konstant

c \* F(x) = ( c\*f0, c\*f1, c\*f2, c\*f3, …)

Additions regeln…

F(x) + G(x) = (f0+g0, f1+g1, f2+g2, f3+g3, …)

Högerskift (Lägga till n st nollor i början av sekvens)

Derivering av sekvens..

d/dx(1 + x + x^2 + …) = d/dx( 1/(1-x) )

1 + 2x + 3x^2 + … = d/dx( 1/(1-x)^2 )

Multiplicera två sekvenser exempel)

(1, 2, 3, 4, …) 🡨🡪 1 / (1-x)^2

(5, 6, 0, 0, …) 🡨🡪 5+6x

(5\*1, 5\*2+6\*1, 5\*3+6\*2+0\*1, 5\*4+6\*3+2\*0’1\*0, …)

Exempel på vanliga sekvenser)

(1, 1, 1, 1, …) 🡨🡪 1 / (1-x)

(1, 2, 3, 4, …) 🡨🡪 d/dx ( 1 / (1-x) ) = 1 / (1-x) ^2

(0, 1, 2, 3, …) 🡨🡪 x / (1-x) ^2

(1, 4, 9, 16, …) 🡨🡪 d(dx ( x / (1-x)^2 ) = (1+x) / (1-x)^3

(0, 1, 4, 9 , …) 🡨🡪 x(1+x) / (1-x)^3

(1, -1, 1, -1, ..) 🡨🡪 1 / (1+x)

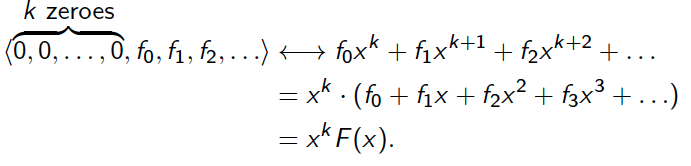
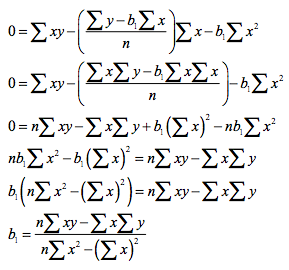
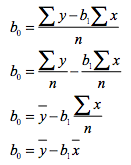
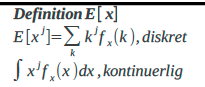
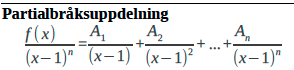
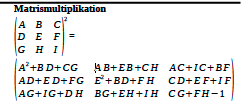
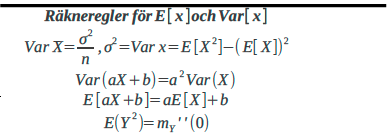
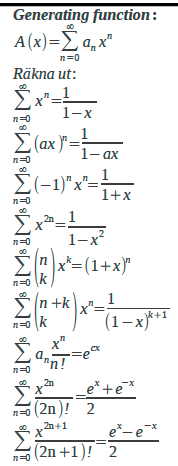
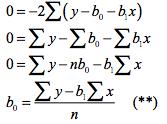
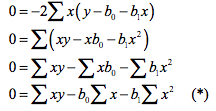
(1, 0, 1, 0, …) 🡨🡪 1 / (1-x^2)

(2, 0, 2, 0, …) 🡨🡪 2 / (1-x^2)

Vad kan man använda generating functions till?

Genererande funktioner är särskilt användbara till att lösa problem av typen ”på hur många sätt kan man välja n st föremål ur ett set”. Koefficienten framför x^n talar om på hur många olika sätt vi kan välja n stycken föremål. Till exempel, den genererande funktionen för binomial koefficienter och genererande funktionen för att välja föremål från k element set med repetition.

Macintosh HD:Users:theodorastrand:Desktop:Skärmavbild 2013-10-19 kl. 16.24.40.png

Macintosh HD:Users:theodorastrand:Desktop:Skärmavbild 2013-10-19 kl. 16.27.47.pngMacintosh HD:Users:theodorastrand:Desktop:Skärmavbild 2013-10-19 kl. 16.32.31.pngMacintosh HD:Users:theodorastrand:Desktop:Skärmavbild 2013-10-19 kl. 16.27.57.png

Simple way to calculate linear regression)

Y(tak) = b0 + b1 X

b0 = y(mean)-b1\*x(mean)

b1 = ( x(mean) \* y(mean) – (x\*y)(mean) ) / ( x(mean)^2 – x^2(mean) )

Exempel)

(7.3, 10.2), (3.2,6.5), (8.7,14), (6.4,9.9), (1.1,3.6), (2.6,5.9)

Hitta least square regression line y= ax + b..

X(mean) = (7.3+3.2+8.7+6.4+1.1+2.6)/6 = 4.8833..

Y(mean) = (10.2+6.5+14+9.9+3.6+5.9)/6 = 8.35

(x\*y)(mean) = (7.3\*10.2 + 3.2\*6.5 + 8.7\*14 + 6.4\*9.9 + 1.1\*3.6 + 2.6\*5.9)/6 = 49.9533..

x^2(mean) = (7.3^2+3.2^2+8.7^2+6.4^2+1.1^2+2.6^2)/6 = 31.35833..

B1 = ( 4.8833 \* 8.35 – 49.9533 ) / ( 4.8833^2 – 31.3583 ) = 1.2218

B0 = 8.35 – 1.2218 \*4.8833 = 2.38358406

SSE := (10.2-2.38358406-1.2218\*7.3)^2+(6.5-2.38358406-1.2218\*3.2)^2+(14-2.38358406-1.2218\*8.7)^2+(9.9-2.38358406-1.2218\*6.4)^2+(3.6-2.38358406-1.2218\*1.1)^2+(5.9-2.38358406-1.2218\*2.6)^2 = 2.455959

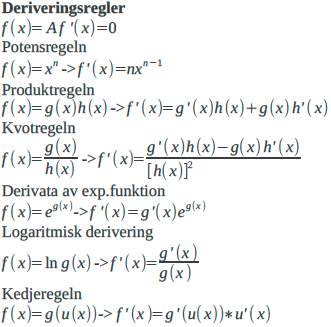
**SSR =**

**SST = SSR + SSE**

Vad används linear regressison till?

If the goal is prediction, or forecasting, or reduction, linear regression can be used to fit a predictive model to an observed data set of*y* and *X* values. After developing such a model, if an additional value of *X* is then given without its accompanying value of *y*, the fitted model can be used to make a prediction of the value of *y*.

Given a variable *y* and a number of variables *X*1, ..., *Xp* that may be related to *y*, linear regression analysis can be applied to quantify the strength of the relationship between *y* and the *Xj*, to assess which *Xj* may have no relationship with *y* at all, and to identify which subsets of the *Xj* contain redundant information about *y*.



Generellt fall för MLE:

Multiplicera ihop alla funktioner.

Ta log av de multiplicerade funktionerna.

Derivera och sätt = 0

Lös ut variabel

f(x1,x2,x3...|a) = f(x1|a) \* f(x2|a)\*... = lik(a)

lik(a) = ∏ f(xi|a)

ln(lik(a)) = l(a) = sum(ln(f(x|a))

d/dx(ln(lik(a))) = 0

Fördelar mellan mm och mle:

Mle är att föredra över mm i de flesta fall eftersom att mle maximerar sannolikheten att de stämmer och att mm tar ett medelvärde från en del av populationen och behöver därav inte vara speciellt precis.

Dock ger MLE och MM oftas samma resultat.

Logaritm lagar (kan användas vid MLE)

10x = y 🡨🡪 x = lg y

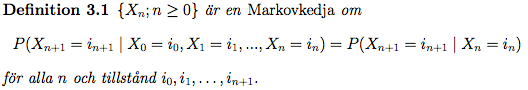
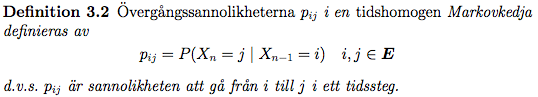
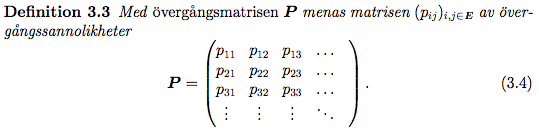
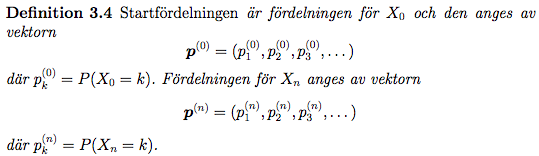
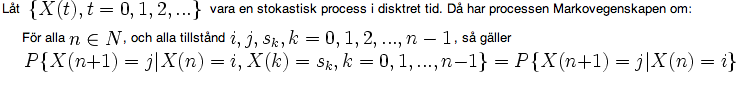
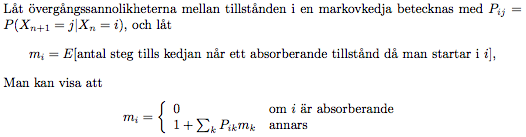
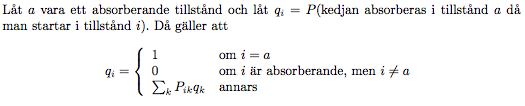
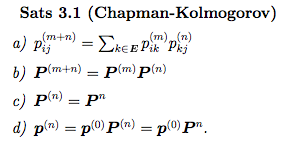
ex = y 🡨 🡪 x = ln y

lg xy = lg x + lg y

lg (x/y) = lg x – lg y

lg xp = p lg x

1 / (1-x)^k = **∑ (k+n-1n) xn**



Uppgift 4)

Let *X*1, *X*2, ..., *Xn* be Bernoulli random variables with parameter *p*. What is the method of moments estimator of *p*?

E(Xi) = p

We have just one parameter for which we are trying to derive the method of moments estimator.  Therefore, we need just one equation. Equating the first theoretical moment about the origin with the corresponding sample moment, we get: (SUMMA FORMELN GÅR MELLAN i=1 och n)

p= (1/n) \* ∑ Xi

Gör om p till en estimator (sätt en hatt på p):

^p = (1/n) \* ∑ Xi

So, in this case, the method of moments estimator is the same as the maximum likelihood estimator, namely, the sample proportion.

Let *X*1, *X*2, ..., *Xn* be normal random variables with mean *μ* and variance *σ*2. What are the method of moments estimators of the mean *μ* and variance *σ*2?

*E*(*Xi*) = *μ*and*E*(*Xi*2) =*σ*2+ *μ*2

E(x) = *μ* = (1/n) \* ∑ Xi (SUMMA FORMELN GÅR MELLAN i=1 och n)

E(X2) = σ2 + μ2 = (1/n) \* ∑ Xi2

Now, the first equation tells us that the method of moments estimator for the mean *μ* is the sample mean:

μ^MM= (1/n) \* ∑ Xi = X(mean)

And, substituting the sample mean in for *μ* in the second equation and solving for *σ*2, we get that the method of moments estimator for the variance *σ*2is:

*σ*2MM = (1/n) \* ∑ Xi2 - μ2 = (1/n) \* ∑ Xi2 – X2(mean)

which can be rewritten as:

*σ*2MM = (1/n) \* ∑ (Xi - X(mean))2

Again, for this example, the method of moments estimators are the same as the maximum likelihood estimators.

Suppose that X follows a geometric distribution, P(X = K) = p(1-p)k-1 and assume an iid sample of size n. Find the method of moments estimator for p: (i = 1 och övregräns oändligheten)

E(X) = ∑ kp(1-p)k-1 = p \* ∑ k(1-p)k-1 = p / (1-(1-p))2 = 1/p

So the MME estimator of p is ^p = 1 / x(mean)

Let X1, . . . , Xn be iid Binomial(n, p), that is,

P(Xi = x | n, p) = (nx) px (1-p)n-x, x = 0, 1, . . . , n.

Here we assume that both k and p are unknown and we desire point estimators for both parameters.

X(mean) = np

(1/n) \* ∑ Xi2 = np(1-p) + n2 \* p2 , Solving for n and p yields the estimators..

n = (x(mean))2 / x(mean) – (1/n) ∑ (Xi– X(mean))2 and P = X(mean) / n

method of moments poisson; To estimate parameter λ of Poisson(λ) distribution, we recall that μ1 = E(X) = λ

There is only one unknown parameter, hence we write one equation,

μ1 = λ = m1 = X(mean)

”Solving” it for λ, we obtain λ(mean) = X(mean),

the method of moments estimator of λ.

MLE geometric; Let *X*1,*X*2,*X*3.....*Xn* be a random sample from the geometric distribution with p.d.f.

https://www.projectrhea.org/rhea/images/math/e/9/2/e924df0b8fd59eac6f1543d2cfd3855c.png The likelihood function is given by:

taking log,

https://www.projectrhea.org/rhea/images/math/8/1/7/817251a9f57176d9e79eb335c0f9553a.png Differentiating and equating to zero, we get,

https://www.projectrhea.org/rhea/images/math/3/f/e/3fe50ccfd665f407db375da22c748e5c.png THEREFORE,

https://www.projectrhea.org/rhea/images/math/e/3/9/e39c8d10b3a1ee2292e6b37ada910fed.png So, the maximum likelihood estimator of P is:

https://www.projectrhea.org/rhea/images/math/9/a/b/9ab0572c79f2db9eae981807048ed532.png

MLE BERNOULLI: Since *X*1,*X*2, . . . ,*Xn* are iid random variables, the joint distribution is

L(p; x)≈f(x; p)=∏f(xi;p)=∏px(1−p)1−x Differentiating the log of L(p ; x) with respect to p and setting the derivative to zero shows that this function achieves a maximum at

 p^= (1/n)∑xi

https://www.projectrhea.org/rhea/images/math/b/e/a/beaccb61948a46cc3747ba394877f68c.png

Binomial PMF:



Bernoulli PMF:

Geometric PMF:

Possion PMF:

Normal PDF:



Låt n vara det totala antalet aborrar i sjön. Visa att ^n = 10x är en väntevärdig skattare för n.

X~Geom(10/n) ger E[^n] = E[10x] = 10 E[x] = 10 \*n/10 = n

Vad är Var [^n]?

Var ^n = Var [10x] = 100 Var [x] =

If X1, .. Xn are iid Normal(*μ ,* σ2), their joint density is the product of their marginal densities:

F(X1, .. Xn | *μ ,* σ) =

∏ 1 / (σ \*sqrt(2\*pi)) exp ( -1/2 [(Xi – μ)/ σ]2

Regarded as a function of *μ and* σ, this is the likelihood function . The log likelihood is thus;

L(*μ ,* σ) =

-n \* log σ – n/2 \* log 2pi – 1/2σ2 ∑ (Xi - *μ*)2

The partials with respect to *μ and* σ are:

dL/dμ => 1/ σ2 ∑ (Xi - *μ*)

dL/dσ => - n/σ + σ-3 ∑ (Xi - *μ*)2

Setting the first partial to zero and solve for mle gives; *μ(TAK) = x(mean)*

*Setting the second partial to zero and substituting the mle for μ we find the mle:*

Gives; σ(TAK) = sqrt( (1/n) \* ∑ (Xi – *X(mean)*)2

Uppgift 1)

Markov egenskapen defineras;  
  
  
  
  
Definition för markovkedja..

If X follows poisson distribution with parameter λ then

P(X = x) = λx e-λ / x! If x1, … xn are iid and poisson their joint frequency function is the product of the marginal frequency functions. The log likelihood is thus;

L(λ) = ∑ ( Xi log λ – λ – log Xi!)

=log λ ∑ Xi – nλ – ∑ log Xi!

Settin the first derivate of the log likelihoog equal to zero we find:

L’(λ) = (1/ λ) ∑ Xi - n = 0

The mle is then;

Lambda(TAK) = x(mean)

MLE binomial: Suppose that X is an observation from a binomial distribution, X ~Bin(n, p), where n is known and p is to be estimated. The likelihood function is:

L(p; x) = n! / (x! (n-x)! ) \* px (1-p)n-x which, except for the factor n! / (x! (n-x)! ) is identical to the likelihood from n independent bernoulli trials with p= (1/n) \* ∑ Xi . But since likelihood is regarded as a function only of parameter p the factor is just a fixed constant and does not affect the MLE. Therefore,

\_\_\_\_\_\_\_\_\_\_, not affect the MLE. ant and does not ted. the asdo asod oasdo asdo asajnd ajdns ajsdn asjn aamdksadmaskdmsaksdmas p = x/n