Formule PS

|A ∪ B| = |A| + |B| − |A ∩ B|

Repartitia (distributia) Poisson:

Daca L si R sunt disjuncte: P(L ∪ R) = P(L) + P(R)

Daca L si R nu sunt disjuncte: P(L ∪ R) = P(L) + P(R) − P(L ∩ R) (principiul includerii si excluderii)

Probabilitatea conditionata:

Regula multiplicarii: P(A ∩ B) = P(A|B) · P(B)

Legea probabilitatii totale (B1, B2, B3 disjuncte 2 cate 2): P(A) = P(A ∩ B1) + P(A ∩ B2) + P(A ∩ B3)

P(A) = P(A|B1)P(B1) + P(A|B2)P(B2) + P(A|B3)P(B3)

Daca A e independent de B: P(A|B) = P(A)

Formula lui Bayes:

Repartitia binomiala:

Repartitia geometrica:

Media (v.a. discreta):

Media unei variabile aleatoare Bernoulli(p) este p.

X ~ binomial(n, p):



Media unei repartitii geometrice:

Variatia (dispersia): , unde

Deviatia standard:



Daca X ~ Bernoulli(p):

X si Y sunt v.a. independente ⬄ P(X = a, Y = b) = P(X = a)P(Y = b), ∀a, b

X si Y sunt v.a. independente ⇒ Var(X+Y ) = Var(X)+Var(Y )

Var(aX + b) = a^2Var(X)

Var(X) = E(X^2) − E(X)^2

X ~ binomial(n, p): Var(X) = np(1 − p)

A math equations and formulas on a white background

Description automatically generated

X v.a. continua: , f se numeste pdf (functia densitate de probabilitate)

Proprietati pdf:

1. f(x) ≥ 0 (f este nenegativa)
2. 

X v.a. continua: , f se numeste cdf (functia de distributie cumulativa)

sau functia de repartitie

Proprietati cdf:

1. F(x) = P(X ≤ x) (definitia)
2. 0 ≤ F(X) ≤ 1, ∀x ∈ R.
3. F este crescatoare, i.e. daca a ≤ b, atunci F(a) ≤ F(b)
4. P(a ≤ X ≤ b) = F(b) − F(a)
5. F′(x) = f(x), ∀x in care F este derivabila

Media (v.a. continua):

h functie continua, atunci Y = h(X) e v.a. si

A math equations and graphs

Description automatically generated with medium confidence

A math equations on a white background

Description automatically generated

A comparison of a function

Description automatically generated with medium confidence

A text on a white background

Description automatically generated

A diagram of a normal distribution

Description automatically generated

A white background with black text

Description automatically generated

Mediana lui X este valoarea x pentru care P(X ≤ x) = 0.5, i.e. valoarea lui x astfel incat P(X ≤ x) = P(X ≥ x). Cu alte cuvinte, X are probabilitate egala de a fi deasupra sau sub mediana ¸si de aceea fiecare probabilitate este 1/2. In termeni de cdf F(x) = P(X ≤ x), putem defini echivalent mediana ca valoarea lui x satisfacand F(x) = 0.5.

A p-a quantila a lui X este valoarea qp astfel incat P(X ≤ qp) = p.

Legea numerelor mari (LoLN):

1. Media multor date independente este (cu mare probabilitate) aproape de media repartitiei de baza.
2. 2. Histograma densitatii multor date independente este (cu mare probabilitate) aproape de graficul densitatii repartitiei de baza.

Teorema limita centrala (CLT): Suma sau media multor copii independente ale unei variabile aleatoare este aproximativ o variabila aleatoare normala.

X1,X2, ...,Xn sunt variabile aleatoare independente cu aceeasi repartitie ⇒ Xi sunt independente si identic distribuite (i.i.d.). Xi au aceeasi medie μ si deviatie standard σ.

Media lui X1, X2, ...,Xn:

LoLN: Cand n creste, probabilitatea ca Xn este aproape de μ tinde la 1.

CLT: Cand n creste, repartitia lui Xn converge la repartitia normala N(μ, σ2/n).

LoLN (enunt formal): Presupunem ca X1,X2, ...,Xn, ... sunt variabile aleatoare i.i.d. cu media μ si dispersia σ^2. Pentru fiecare n, fie Xn media primelor n variabile. Atunci pentru orice a > 0, avem

LoLN pentru histograme: Cu probabilitate mare histograma de densitate a unui mare numar de date dintr-o repartitie este o buna aproximare a graficului pdf f a repartitiei.

Fiind data o variabila aleatoare X cu media μ si deviatia standard σ, definim standardizarea lui X ca noua variabila aleatoare

Z are media 0 si deviatia standard 1. Daca X are o repartitie normala, atunci standardizarea lui X are repartitia normala standard Z cu media 0 si dispersia 1.

CLT (enunt): Presupunem ca X1,X2, ...,Xn, ... sunt variabile aleatoare i.i.d., fiecare avand media μ si deviatia standard σ. Pentru fiecare n notam cu Sn suma si cu Xn(barat) media lui X1, ...,Xn.

A mathematical equation with black text

Description automatically generated

Deoarece Sn este multiplu al lui Xn(barat), Sn si Xn(barat) au aceeasi standardizare:



CLT: pentru n mare,

Daca Z ∼ N(0, 1), atunci cu rotunjire avem:

1. P(|Z| < 1) = 0.68
2. P(|Z| < 2) = 0.95; mai precis, P(|Z| < 1.96) ≈ 0.95
3. P(|Z| < 3) = 0.997

Repartitia comuna (v.a. discrete):

1. 0 ≤ p(xi, yj) ≤ 1, ∀i = 1,n, j = 1,m
2. 



cdf comuna:

Repartitia comuna (v.a. continue):

1. f(x, y) ≥ 0, ∀x, y
2. 



cdf comuna: ⇒

cdf comuna F(x, y) a lui X si Y trebuie sa satisfaca proprietatile:

1. F este crescatoare: i.e. daca x sau y cresc, atunci F trebuie sa ramana constanta sau sa creasca
2. F(x, y) = 0 la stanga jos a domeniului de valori comun
3. F(x, y) = 1 la dreapta sus a domeniului de valori comun

V.a. repartizate comun X si Y sunt independente ⬄ cdf comuna a lor e produsul cdf-urilor marginale: F(x, y) = FX(x)FY (y)

Pt v.a. discrete, echivalent cu: p(xi, yj) = pX(xi)pY (yj)

Pt v.a. continue, echivalent cu: f(x, y) = fX(x)fY (y)

Covarianta: Cov(X, Y) = E ((X − μX) (Y − μY )) , unde X si Y sunt variabile aleatoare cu mediile μX si μY

Proprietati covarianta:

1. Cov(aX + b, cY + d) = acCov(X, Y), ∀a, b, c, d constante
2. Cov(X1 + X2, Y ) = Cov(X1, Y) + Cov(X2, Y)
3. Cov(X,X) = Var(X)
4. Cov(X, Y) = E(XY) − μXμY
5. Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y), ∀X, Y
6. Daca X si Y sunt independente, atunci Cov(X, Y) = 0 (reciproca e falsa)

X, Y v.a. discrete, pmf comuna p(xi, yj):



X, Y v.a. continue, pmf comuna p(xi, yj):



Coeficientul de corelatie intre X si Y este definit de

Proprietati corelatie:

1. ρ este covarianta standardizarilor lui X ¸si Y
2. ρ este fara dimensiune
3. −1 ≤ ρ ≤ 1. Mai mult,

ρ = 1 ⬄ Y = aX + b cu a > 0

ρ = −1 ⬄ Y = aX + b cu a < 0