Matematisk statistik TAMS27 formelblad

Mattias Salo

27 maj 2018

Väntevärde:
Kontinuerlig

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx$$

Diskret (Ω är alla möjliga k)

$$E[X] = \sum_{k\in\Omega} kP(X = k)$$

Varians:
Kontinuerlig

$$Var(X) = V(X) = \int_{-\infty}^{\infty} (x - E[X])^2 f_X(x)dx$$

förkortningsformel:

$$Var(X) = E[X^2] - (E[X])^2 = \int_{-\infty}^{\infty} x^2 f_X(x)dx - (E[X])^2$$

Diskret

$$Var(X) = V(X) = \sum_{k\in\Omega} (k - E[X]^2 P(X = k))$$

förkortningsformel:

$$Var(X) = E[X^2] - (E[X])^2 = \sum_k k^2 P(X = k) - (E[X])^2$$

Standardavvikelse:

$$D(X) = \sqrt{Var(x)}$$

Sats Räkneregler: (*Diskret*)
a) $E[a] = a, Var(a) = 0$ för varje icke slumpmässigt $a \in \mathbb{R}$
b) $E[aX + b] = a * E[X] + b$, $a, b \in \mathbb{R}$
$Var(aX + b) = a^2 Var(X)$ (inget b på högra sidan)
c) Låt g vara en reell funktion. Det gäller att $E[g(X)] = \sum_k g(k)P(X = k)$

Sats Räkneregler: (*Kontinuerlig*)
a) och b) som i det diskreta fallet
c)Låt g vara en reell funktion. Det gäller att $E[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x)dx$.

Simultana fördelningar:
Låt X, Y vara två kontinuerliga slumpvariabler.
Definition:
a) $p(x, y) = P(X = x, Y = y)$ kallas simultan sannolikhetsfunktion för (X,Y).
b)

$$p_X(x) = P(X = x) = \sum_{y:p(x,y)>0} p(x, y)$$

$$p_Y(y) = P(Y = y) = \sum_{x:p(x,y)>0} p(x, y)$$

kallas de marginella sannolikhetsfunktionerna.
Anmärkning:
Den marginella sannolikhetsfunktionen $p_X(x) = P(X = x)$ är samtidigt sannolikhetsfunktionen av s.v. X om man inte tar hänsyn till Y.

Låt X, Y vara två kontinuerliga slumpvariabler.
Definition:
a) en funktion $f : \mathbb{R}^2 \longrightarrow [0, \infty)$ kallas simultan täthetsfunktion av (X, Y) om

$$P((X, Y) \in C) = \iint_{(x,y)\in C} f(x, y)dxdy$$

för varje $C \subseteq \mathbb{R}^2$.
b)

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy, x \in \mathbb{R},$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)dx, y \in \mathbb{R},$$

kallas de marginella täthetsfunktionerna.
Anmärkningar:
(1) den marginella täthetsfunktionen $f_X(x)$ är samtidigt täthetsfunktionen av s.v. X om man inte tar hänsyn till Y.
(2) Det gäller att

$$P(X \in A) = \int_{X\in A} \int_{y=-\infty}^{\infty} f(x, y)dydx = \int_{x\in A} f_X(x)dx, A \subseteq \mathbb{R}.$$

Oberoende slumpvariabler
definition
Låt X, Y vara två s.v. X, Y kallas oberoende om:

$$P(X \in A, Y \in B) = P(X \in A) * P(Y \in B)$$

för vilka två mängder som helst $A, B \subseteq \mathbb{R}$.
Sats: a) Låt X, Y vara två diskreta s.v.
X, Y oberoende om och endast om

$$P(X = i, Y = j) = P(X = i) * P(Y = j)$$

för alla par (i, j) som (X, Y) kan anta. Detta skrivs också som $p(i, j) = p_X(i) * p_Y(j)$för alla par (i, j).
b) Låt X, Y vara två kontinuerliga s.v.
X, Y är oberoende om och endast om

$$f(x, y) = f_X(x) * f_Y(y)$$

för alla $(x, y) \in \mathbb{R}^2$.

Sats: a) Låt $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$. Det gäller att

$$E[g(X, Y)] = \begin{cases} \sum_{alla} \sum_{(k,j)} g(k, j)p(k, j), X, Y diskreta \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)f(x, y)dxdy, X, Y kontinuerliga \end{cases}$$

där $p(k, j)$ är den gemensamma sannolikhetsfunktionen (diskret fall) och $f(x, y)$ är den gemensamma täthetsfunktionen (kontinuerligt fall).
b) Låt X, Y vara oberoende och $g : \mathbb{R} \longrightarrow \mathbb{R}$, $h : \mathbb{R} \longrightarrow \mathbb{R}$.
Det gäller att: $E[g(X) * h(Y)] = E[g(X)] * E[h(Y)]$.
I synnerhet: $E[X * Y] = E[X] * E[Y]$.

Definition
$Cov(X, Y) = C(X, Y) = E[(X - E[X]) * (Y - E[Y])]$.
kallas kovarians av X och Y.
Anmärkningar: (1) Fördelaktigt: Förkortningsformeln:
$Cov(X, Y) = E[X * Y] - E[X] * E[Y]$
(2) I synnerhet, om X och Y är oberoende $Cov(X, Y) = 0$.
(3) Motriktningen kan vara fel.

Sats (räkneregler för Cov)
(1) $Cov(X, X) = Var(X)$
(2) $Cov(X, Y) = Cov(Y, X)$
(3) $Cov(aX, Y) = Cov(X, aY) = a * Cov(X, Y), a \in \mathbb{R}$
(4) $Cov(\sum_{i=1}^m X_i, \sum_{j=1}^n Y_j) = \sum_{i=1}^m \sum_{j=1}^n Cov(X_i, Y_j)$
Följdsats:
(1) $Cov(a, X) = Cov(X, a) = 0, a \in \mathbb{R}$
(2) $Cov(\sum_{i=1}^m a_i X_i + b, \sum_{j=1}^n c_j Y_j + d) = \sum_{i=1}^m \sum_{j=1}^n a_i c_j Cov(X_i, Y_j)$
(3) $Var(\sum_{i=1}^m X_i) = \sum_{i=1}^m Var(X_i) + \sum_{i\neq j} Cov(X_i, X_j)$

Definition:

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X) * Var(Y)}}$$

kallas korrelation mellan X och Y.
Sats a) $|\rho(X, Y)| \leq Var(X)Var(Y)$
b) $|\rho(X, Y)| = 1$ om och endast om det finns $a, b, c \in \mathbb{R}$ sådana att $aX + BY = c$.
Anmärkning a) är ekvivalent med
$(Cov(X, Y))^2 \leq Var(X) * Var(Y)$
den stokastiska versionen av Schwarz' olikhet.

Centrala gränsvärdessatsen:
Låt $X_1, X_2, ...$" vara en oändlig följd av oberoende och likafördelade stokastiska variabler med väntevärde μ och med standardavvikelsen $\sigma > 0$. Låt den stokastiska variabeln

$$Y_n = X_1 + X_2 + ... + X_n$$

beteckna summan av de första n stokastiska variablerna i följden.
Då gäller att

$$\lim_{n\to\infty} P\left(a < \frac{Y_n - n\mu}{\sigma\sqrt{n}} \leq b\right) = \Phi(b) - \Phi(a)$$

där $\Phi(a)$ betecknar fördelningsfunktionen för en standardiserad normalfördelning.

Normalfördelning:
Normalfördelningen har täthetsfunktionen

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

där μ och σ är normalfördelningens karakteristiska konstanter: μ är väntevärdet och σ är standardavvikelsen för fördelningen. Denna normalfördelning betecknas med $N(\mu, \sigma)$.
Normalfördelningens täthetsfunktion kan inte integreras med vanliga endimensionella metoder eftersom den inte har någon primitiv funktion som kan uttryckas analytiskt. Arean under kurvan kan emellertid med andra metoder visas vara 1, vilket den måste vara för att vara en sannolikhetsfördelning.
En **standardiserad normalfördelning** har $\mu = 0$ och $\sigma = 1$.
Fördelningsfunktionen för en standardiserad normalfördelning brukar betecknas med Φ, och sambandet mellan fördelningsfunktion och täthetsfunktion är

$$\Phi(x) = \int_{-\infty}^{x} f(z)dz.$$

Fördelningsfunktionen anger sannolikheten för att en normalfördelad variabel Y är mindre än eller lika med ett visst tal a:

$$P(Y < a) = \Phi(a).$$

Sannolikheten att en normalfördelad variabel Y hamnar i ett intervall $[a, b]$ är

$$P(a < Y < b) = \Phi(b) - \Phi(a).$$

Fördelningsfunktion
Fördelningsfunktionen för en godtycklig normalfördelad variabel $X \in N(\mu, \sigma)$ kan erhållas från fördelningsfunktionen för en standardnormalfördelad variabel:

$$P(X < a) = \Phi\left(\frac{a-\mu}{\sigma}\right).$$

$$P(a < X < b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

Denna egenskap medför att tabeller för normalfördelningar bara redovisar fördelningsfunktionen Φ, eftersom alla andra normalfördelningar på detta sätt kan översättas till en normalfördelning med väntevärdet 0 och standardavvikelsen 1.
Sats: Om $X_1, ..., X_n$ är oberoende $N(\mu, \sigma)$ och $\overline{X} = \sum_1^n X_i/n$ är deras aritmetiska medelvärde, så gäller att: $\overline{X} \in N(\mu, \sigma/\sqrt{n})$.

Markovkedjor:
Irreducibel är att alla noder går att nå från alla. Aperiodisk är när en nod kan gå till sig själv. Om den både är irreducibel och aperiodisk är den ergodisk.

Lagen om total sannolikhet:
Det då jag skriver ett träd är lagen om total sannolikhet.
Bayes' sats:
Låt $A_1, ..., A_n$ vara n disjunkta (oförenliga) händelser med positiv sannolikhet. Anta att Händelserna utgör hela utfallsrummet $\cup_{i=1}^n A_i = \Omega$. Bayes' sats säger då att

$$P(A_i|B) = \frac{P(Ai)P(B|A_i)}{\sum_{j=1}^n P(A_j)P(B|A_j)}$$

där nämnaren är lika med $P(B)$ enligt lagen om total sannolikhet.
För specialfallet $n = 1$ ger Bayes sats

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$

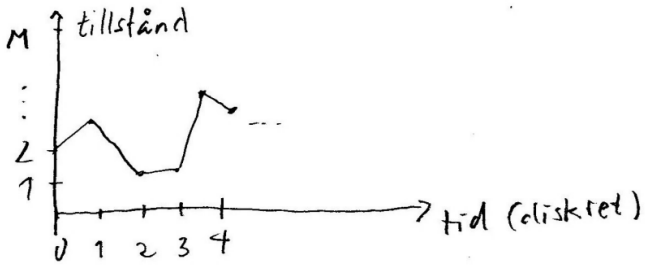
Där $P(A|B)$ är sannolikheten för A, givet B.

Betingad sannolikhet

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Om en Poissonfördelning används och du skall räkna ut en sannolikhet över flera händelser kan du multiplicera antalet händelser med det originella μ-värdet.
$X \sim Po(\mu)$ och $Y = X_1, X_2, ..., X_n$ kan det sägas att $Y \sim Po(n * \mu)$.

Markovkedjor
definition: a) Låt $X_0, X_1, ...$ vara en följd av stokastiska variabler som kan anta värdena $0, 1, 2, ..., M$. Man säger att systemet är vid tid n i tillståndet $i \in 0, 1, ..., M$ om $X_n = i$.



b) Följden $X_0, X_1, ...$ bildar en Markovkedja om varje gång när systemet är i tillstånd $i \in 0, 1, ..., M$ finns det en sannolikhet p_{ij} att systemet kommer att vara i tillstånd j näst. Dvs för alla $i_0, ..., i_{n-1}, i, j$.

$$p_{ij} = P(X_{n+1} = j|X_n = i, X_{n-1} = i_{n-1}, ..., X_0 = i_0)$$

c) Värdena p_{ij}, $0 \leq i, j \leq M$, kallas övergångssannolikhet. Matrisen $p = [p_{ij}]_{ij} = 0, ..., M$ kallas övergångsmatris.
Anmärkning $P(X_{n+1} = j|X_n = i, X_{n-1} = i_{n-1}, ..., X_0 = i_0) = p_{ij} = P(X_{n+n} = j|X_n = i)$ Markovprincipen.

Sats (Chapman-Kolmogorov ekvationer), Låt $p_{ij}^{(n)}$ vara sannolikheten att systemet är i tillstånd i och kommer att vara i tillstånd j efter n steg (tidsenheter). Det gäller att

$$p_{ij}^{(n)} = \sum_{k=0}^{M} p_{ik}^{(r)} * p_{kj}^{(n-r)}$$

för alla $0 < r < n$.
Anmärkningar 1) Låt

$$P^{(n)} = [p_{ij}^{(n)}]_{i,j=0,...,M}$$

beteckna n-steg övergångsmatrisen
Satsen kan skrivas som

$$P^{(n)} = p^{(r)} * p^{(n-r)}$$

(matrismultiplikation)
2) I synnerhet $p = p^{(1)}$ och

$$p^{(n)} = p * .. * p = p^n$$

(matrispotens)