Probabilitat I Var. Aleatòries (A)

Experiència aleatòria. Probabilitat

Experiència aleatoria. Definicions.

- Fenòmen Determinista: Mateixos resultats a partir mateixes cond. inicials.
- Fenòmen Aleatòri: Incertesa en el resultat en exp. aleatòria.
- Conjut Resultats possibles: Associat a una exp. aleatòria. $\Omega = \{1,2,3,4,5,6\}$ #Dau
- Esdeveniment o Succés: Subconjut del Conj. Res. Possibles. $A_{parr} \subseteq \Omega$; $A_{senars} \subseteq \Omega$
- Partició: Conjut esdeveniments $A_i \neq 0 \wedge$ disjunts t.q. $A_1 \cup A_2 \cup \dots \cup A_k = \Omega$. $A_p \cap A_s = \emptyset$ · $A_p \cup A_s = \Omega$

Probabilitat. Definició i Propietats

$\boxed{0 \leq P(A) \leq 1}$ $\boxed{P(A_1 \cup \dots \cup A_n) = P(A_1) + \dots + P(A_n)}$ $\boxed{P(A \cup B) = P(A) + P(B) - P(A \cup B)}$

! Laplase no funciona si no són totes particions equiprobables.

Independència

Es una propietat de les probabilitats. $\boxed{P(A) \text{ i } P(B) \text{ indep.} \Leftrightarrow P(A \cap B) = P(A) \cdot P(B)}$

#Cas de dubte no ho son

Probabilitat Condicionada

$P(A|B) \equiv$ "Probabilitat de A si ha succeit B".

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

NO és prob. indep

Limitar al conj. observat

Probabilitat a Posteriori

$P(B_k|A) \equiv$ "Quina és la prob. de que hagi passat B_k si ha passat A".

$$\boxed{P(B_k|A) = \frac{P(B_k \cap A)}{\sum_{i=1}^{n} [P(B_i) \cdot P(B_i|A)]}}$$

Prob d'aquest B_k concret i A dividit entre totes les possibles B_i.

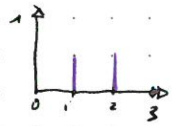
Variable Aleatoria (VA)

- Discreta (VAD): Resultat de la variable està dins d'un conj. enumerable.
- Continua (VAC): Pot tindre qualsevol valor $\mathbb{R}$ en un interval.

Funció de Prob. i dist. de VAD

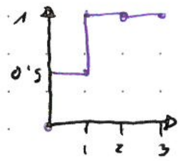
- Funció de prob ($P_X(k)$): Probabilitat puntual de cadascun dels possibles k.

$\boxed{P_X(k) = P(X=k)}$. Recordem que ha de complir $\sum_{k=1}^{m} P_X(k) = 1$.



- Funció de distribució ($F_X(k)$): Probabilitat acumulada d'una VAD.

$$\boxed{F_X(k) = P(X \le k) = \sum_{j \le k} P_k(j)}$$



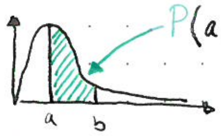
Funció de Prob i dist. de VAC

- Funció de densitat ($f_Y(x)$): Funció (positiva) que indica quant de probable és que la Y estigui aprop de x. Modela les dades.

Pq. una funció ho sigui, ha de complir: $\boxed{f_Y(x) \ge 0}$ i $\boxed{\int_{-\infty}^{+\infty} f_Y(x)\,dx = 1}$

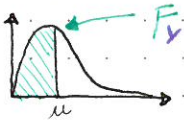
NO diu la prob de Y en un punt x, sino que hem de treballar en intervals (a,b).

$$\boxed{\int_a^b f_Y(x)\,dx = F(b) - F(a) = P(a \le Y \le b)}$$



- Funció de distribució ($F_Y(u)$): Probabilitat acumulada en VAC.

$$\boxed{F_Y(u) = \int_{-\infty}^{u} f_Y(x)\,dx = P(Y \le u)}$$



Fixe't que no fem servir k, sino u. Això es per diferenciar que k és un élement específic dins del conj. d'opcions. Mentre que u és un $\mathbb{R}$.

Probabilitats Cond. i Indep.

- Cond: $\boxed{P(Z \le a \mid Z' \le b) = P(Z \le a \cap Z' \le b) \,/\, P(Z' \le b)}$
- Indep: $\boxed{P(Z \le a \cap Z' \le b) = P(Z \le a) \cdot P(Z' \le b)}$

Quantils

Mesura estadística que divideix conj. de dades ordenades en parts iguals.

El quantil p és el valor de x_p que compleix $P(Z \le x_p) = p$

"Q2" ≡ "Mitjana"
Es el valor que divideix la prob. amb dues parts iguals
prob. acum. esq. = prob. acum. dreta

Indicadors en Variables Aleatòries

Permeten resumir característiques principals de VA. Son paràmetres, no dades reals.

RECORDA: $\bar{X}$ = Mitjana Mostral. $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ $S_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$

S_x = Desviació Estàndar Mostral. Necessites totes les dades per calcular.

- Valor Esperat = Esperança: Mesura la tendència central. "A quin número tendeix si repetim moltes vegades."
  - □ VAD: $\boxed{E(X)=\mu_X = \sum_{\forall k}(k \cdot p_X(k))}$
  - □ VAC: $\boxed{E(Y)=\mu_Y = \int_{-\infty}^{+\infty} x \cdot f_Y(x)\, dx}$

Podem obs. que tenen una forma similar.

$\sum \rightarrow \int$ i

Om. en un cas serà funció de prob. i l'altre densitat.

- Variança: Medeix la dispersió de la VA respecte el Valor Esperat. "En finances, una variança alta ⇒ major risc." "Si molts treuen 8 i 9 hi haurà poca dispersió, mentre que si van entre 1 a 10, hi haurà molta".
  - □ VAD: $\boxed{Var(X)=\sigma_X^2 = \sum_{\forall k}\left[(k-\mu_X)^2 \cdot p_X(k)\right]} \longrightarrow \sigma_X = \sqrt{\sigma_X^2} = \sqrt{Var(X)}$

    Desviació típica/tipus
  - □ VAC: $\boxed{Var(Y)=\sigma_Y^2 = \int_{-\infty}^{+\infty}(x-\mu_Y)^2 \cdot f_Y(x)\, dx} \longrightarrow \sigma_Y = \sqrt{\sigma_Y^2} = \sqrt{Var(Y)}$

- Relació entre Variança i Valor Esperat: $\boxed{Var(X) = E(X^2) - E(X)^2}$

Parell de Variables Aleatòries

Funcions de prob. VAD

Tenim dues VA X, Y amb casos de la forma (x_i, y_j)

- Funció de prob. conj.: Probabilitat que passin totes dues a l'hora $\boxed{P_{X,Y}(x,y) = P(X=x \cap Y=y)}$
- Funció de prob. cond.: Prob. que passi X si $Y=y$. $\boxed{P_{X|Y=y}(x) = \frac{P_{X,Y}(x,y)}{P_Y(y)}}$

Funcions de prob VAC

- Funció de densitat conj. ($f_{X,Y}(x,y)$): Descriu prob. que X, Y prenguin valors al voltant (x,y).

Aquesta funció es pot obtindre: $\frac{\partial^2 F_{X,Y}(x,y)}{\partial x \, \partial y} = f_{X,Y}(x,y)$.

- Prob indep: $\boxed{f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)}$
- Prob. cond: $\boxed{f_{X|Y=y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}}$

Volum total format sota $f_{X,Y}(x,y) = 1$ (Sino no seria).

$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y) \cdot dx$

Indicadors

Covariança $Cov(X,Y)$

Medeix quant varien dues variables juntes respecte a les seves respectives esperances.

Indica la direcció de la relació entre dos variables i una mesura de la seva magnitut.

Donat que es treballa en diferents unitats, es sensible a les escales i dificulta interpretació.

Es a dir, que podem treballar X en grams i Y en cm pot donar covariança = 170432 i això no sabem si es molt o poc. Però si que podem obs. que $Cov(X,Y)>0 \Rightarrow$ Relació Directa.

Podem llavors interpretar el resultat:

$Cov(X,Y)>0$: Quan X tendeix a estar per sobre $E(X)$, Y també. Relació Directa.

$Cov(X,Y)=0$: No hi ha cap relació. X i Y són independents.

$Cov(X,Y)<0$: Quan X tendeix a estar davall $E(X)$, Y també. Relació Indirecta.

□ VAD: $$\boxed{Cov(X,Y)=\sum_{\forall x}\sum_{\forall y}(x-E(X))\cdot(y-E(Y))\cdot P_{X,Y}(x,y)}$$

Augmentar hores d'estudi (X) fa augmentar notes (Y). Això sig. que hi ha $Cov(X,Y)$ positiva
Relació Directa

□ VAC: $$\boxed{Cov(X,Y)=\iint_{\forall x,y}(x-E(X))\cdot(y-E(Y))\cdot f_{X,Y}(x,y)\cdot dx\,dy}$$

Correlació $Corr(X,Y)$

Mesura estandaritzada de la $Cov(X,Y)$ que té valors entre $[-1,1]$.

Senyala la direcció de la relació i que tant forta és la seva relació.

Millora la interpretació i comparació.

Podem interpretar resultats com:

$\rho_{X,Y}\approx 1$: Relació lineal forta $\rho_{X,Y}=1$: Relació lineal perfecta

Si X, Y independents $\Rightarrow \rho_{X,Y}=0$!! Al revés no.

$$\boxed{Corr(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)}\cdot\sqrt{Var(Y)}}}$$

Propietats

• Esperança: $\boxed{E(X\cdot Y)=E(X)\cdot E(Y) + Cov(X,Y)}$

$E(Z)$ es sensible a l'escale i calcular $E(X\cdot Y)$ no acaba de mostrar com es relacionen.

Es per això que es suma $Cov(X,Y)$ per representar adicional/restant de la relació X,Y.

Al final volem mesurar la tendencia central de (X,X) en conjunt.

• Variança: Volem saber com varien (Dispersió respecte valor esperat) de dues variables.

Calculem com varien per separat, però ens faltarà saber com interactuen entre si.

$$\boxed{\begin{aligned} V(X+Y) &= V(X)+V(Y)+2Cov(X,Y) \\ V(X-Y) &= V(X)+V(Y)-2Cov(X,Y) \end{aligned}}$$

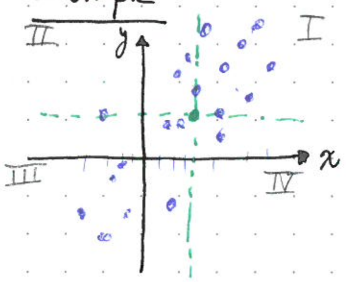
Invertir en dos fondos diferents (X,Y) i saber variança total

Saber variança d'un tractament respecte abans (X) i després (Y)

• Covariança:

$$\boxed{\begin{aligned} Cov(aX,bY) &= a\cdot b\cdot Cov(X,Y) \\ Cov(X,X) &= V(X) \end{aligned}}$$

Exemple



1. Calculem $E(X)$ que és el valor central d'aquesta mostra (respecte X)
2. Calculem $E(Y)$ " " " " " " " (respecte y)
3. Situem $(E(X), E(Y))$ en el pla, serà com el centre gravetat.
4. Calculem la Covariança per saber com varien dues variables conjuntes respecte les seves esperances individuals. Això ens dirà la magnitud.

Si visualment hi ha més punts en I, III $\Rightarrow$ Relació Directa $\Rightarrow Cov(X,Y)>0$
II, IV $\Rightarrow$ Relació Indirecta $\Rightarrow Cov(X,Y)<0$