UNIVERSITAT DE BARCELONA

Esperança condicionada i criteri dels mínims quadrats

Diplomatura d'Estadística

Estadística Matemàtica I

Jordi Ocaña Rebull

Departament d'Estadística

Divisió de Ciències Experimentals i Matemàtiques

UNIVERSITAT DE BARCELONA

Punts que tractarem:

- Concepte d'esperança condicionada. Fórmules de càlcul.
- La variable aleatòria "esperança condicionada". Esperança i variància de l'esperança condicionada.
- Criteri dels mínims quadrats, corba de regressió de la mitjana.
- Regressió lineal mínim quadràtica i correlació.

UNIVERSITAT DE BARCELONA

Concepte d'esperança condicionada

- La distribució condicionada $Y \mid X = x$ descriu les probabilitats associades als possibles valors de Y quan donem per segur que la v.a. X ha pres determinat valor, $X = x$. És lògic que ens preguntem quina és **la seva esperança**: $E(Y \mid X = x)$, la mitjana de tots els possibles valors de Y associats a $X = x$.

Càlcul de l'esperança condicionada

- Cas discret:

$$E(Y \mid X = x) = \sum_j y_j \, f\left(y_j \mid X = x\right)$$

- Cas absolutament continu:

$$E\left(Y \mid X = x\right) = \int_{-\infty}^{+\infty} y \, f(y \mid X = x)\, dy$$

i més en general:

$$E\left(g(Y) \mid X = x\right) = \begin{cases} \sum_j g\left(y_j\right) f\left(y_j \mid X = x\right) \\ \int_{-\infty}^{+\infty} g\left(y\right) f(y \mid X = x)\, dy \end{cases}$$

Altres moments condicionats, variància condicionada

- La darrera fórmula anterior permet definir i calcular qualsevol moment condicionat, com $E(Y^k|X=x)$ i en particular la **variància condicionada** (cas: $g(Y) = (Y - E(Y|X=x))^2$):

$$\operatorname{var}(Y \mid X = x) = \begin{cases} \sum_j \left(y_j - E(Y \mid X = x)\right)^2 f(y_j \mid X = x) \\ \int_{-\infty}^{+\infty} (y - E(Y \mid X = x))^2 f(y \mid X = x)\, dy \end{cases}$$

UNIVERSITAT DE BARCELONA

L'esperança condicionada com a variable aleatòria

- En variar x pot variar $f(y|X=x)$ i per tant $E(Y|X=x)$ (o qualsevol altre moment $E(g(Y)|X=x)$) pot ser diferent: **és una funció de x !!**

$$h(x) = E\,(Y \mid X = x\,)$$

- Per tant, entesa com una funció de la v.a. X, $h(X) = E(\mathrm{Y}|\mathrm{X})$ és una v.a.:

$$\begin{array}{ccccc} \Omega & \xrightarrow{\quad X \quad} & \mathbb{R} & \xrightarrow{\quad E(Y|X=\,) \quad} & \mathbb{R} \\ \omega & \mapsto & X(\omega) & \mapsto & E\,(Y \mid X = X(\omega)) \end{array}$$

UNIVERSITAT DE BARCELONA

Moments de la v.a. esperança condicionada: Teorema fonamental

- La v.a. $E(Y\,|X)$ en realitat "mesura" la mateixa cosa que Y però amb més precisió (amb menys variància):
  1. $E_X\{E(Y \mid X)\} = E(Y)$
  2. $\mathrm{var}_X\{E(Y \mid X)\} \le \mathrm{var}_Y(Y)$

  la desigualtat 2 és conseqüència directa de la important igualtat

$$\mathrm{var}_Y(Y) = \mathrm{var}_X\{E(Y \mid X)\} + E_X\left\{\overbrace{\mathrm{var}(Y \mid X)}^{\ge 0}\right\}$$

UNIVERSITAT DE BARCELONA

Error quadràtic mitjà (MSE)

- Volem predir el valor d'una v.a. Y a partir de certa funció $h(X)$ d'una altra v.a. X (regressió de Y sobre X).
- La v.a. $(Y - h(X))^2$ "error quadràtic" és una forma (no l'única) de mesurar la "discrepància" entre els veritables valors de Y i les prediccions $h(X)$.
- MSE $= E\{(Y - h(X))^2\}$ és una mesura global d'aquesta discrepància.

UNIVERSITAT DE BARCELONA

Criteri dels mínims quadrats i esperança condicionada

- Suposem que en realitat busquem "la millor" funció $h(X)$ per predir Y.
- Criteri dels mínims quadrats: escollir aquella $h(X)$ que fa mínim MSE:

$$\mathrm{MSE}_{\min} = \mathrm{MSE}(h) = \min_{h_i}\{\mathrm{MSE}(h_i)\}$$

- Aquesta funció és sempre l'esperança condicionada: $h(x) = E(Y \mid X = x)$, la **"corba de regressió de la mitjana"**

UNIVERSITAT DE BARCELONA

Raó de correlació

- Corba de regressió de la mitjana: millor regressió de Y sobre X segons mínims quadrats: $\text{MSE}_{\text{min}}=\text{MSE}(E(Y|x))$
- Però, com de bona? La **raó de correlació** mesura aquest concepte:

$$\eta^2(Y \mid X) = 1 - \frac{\text{MSE}_{\text{min}}}{\text{var}(Y)}$$

Gens d'ajust

$$0 \leq \eta^2(Y \mid X) \leq 1$$

Ajust total

UNIVERSITAT DE BARCELONA

Aclariments sobre diapositiva anterior

- Definició i el sentit de raó de correlació: venen de la important igualtat

$$\mathrm{var}(Y) = \underbrace{\mathrm{MSE}_{\min}}_{\text{atribuïble a error}} + \underbrace{\mathrm{var}\{E(Y \mid X)\}}_{\text{atribuïble a regressió}}$$

- Conseqüència d'anterior i de teorema fonamental d'esperança condicionada:

$$E\{\mathrm{var}(E(Y \mid X))\} = MSE_{\min}$$

Regressió lineal mínim quadràtica

- Corba de regressió de la mitjana pot ser intractable, sovint més pràctic funcions més senzilles, p.e. lineals $h(x) = \beta x + \alpha$."
- "Millor" funció lineal?: escollir β i α segons el criteri dels mínims quadrats

$$\frac{\partial\, \mathrm{MSE}(\alpha, \beta)}{\partial \alpha} = 0 \qquad \frac{\partial\, \mathrm{MSE}(\alpha, \beta)}{\partial \beta} = 0$$

$$\Rightarrow \beta = \frac{\mathrm{cov}(X, Y)}{\mathrm{var}(X)} \qquad \alpha = E(Y) - \beta\, E(X)$$

Regressió lineal i coeficient de correlació

- Els anteriors α i β minimitzen l'error quadràtic mitjà en l'àmbit de les funcions lineals: $\mathrm{MSE}_{\min}(\alpha, \beta)$.
- La variància es pot descomposar en

$$\mathrm{var}(Y) = \mathrm{MSE}_{\min}(\alpha, \beta) + \beta^2 \,\mathrm{var}(X)$$

- i fent operacions:

$$\mathrm{MSE}_{\min}(\alpha, \beta) = \mathrm{var}(Y)\big(1 - \rho^2(X, Y)\big)$$

UNIVERSITAT DE BARCELONA

Regressió lineal per mínims quadrats i coef. de determinació

- De les darreres expressions:
  - Si $\rho = 0$ també $\rho^2 = 0$: tota la variància de Y com error: la regressió no explica res
  $$\text{MSE}_{\text{min}}(\alpha, \beta) = \text{var}(Y)$$
  - Si $\rho = \pm 1$ també $\rho^2 = 1$: tota la variància de Y és funció de var(X): regressió lineal total
  $$\text{MSE}_{\text{min}}(\alpha, \beta) = 0 \Rightarrow \text{var}(Y) = \beta^2\, \text{var}(X)$$
  - Per valors intermitjos de ρ^2 (**coeficient de determinació**) $100\rho^2$ és el % d'ajust lineal mínim quadràtic.