CMSC 25025 / STAT 37601

Machine Learning & Large Scale Data Analysis

Kevin Song

4/6/2017.

Problem 1a.

Find density of Y. X have a continuous, strictly increasing cdf F.

$Y = F(X)$.

$$F_Y(x) = Pr\{Y \le x\} = Pr\{F(X) \le x\} = Pr\{X \le F^{-1}(x)\}$$

$$= F(F^{-1}(x)) = \underline{\underline{x}}$$

Now let $U \sim Uniform(0,1)$ & let $X = F^{-1}(U)$. Show that $X \sim F$

$$F_X(x) = Pr\{X \le x\} = Pr\{F^{-1}(U) \le x\} = Pr\{U \le F(x)\}$$

$$= F(x)$$

$\therefore X \sim F$

Problem 1b

Let $X, Y \sim$ Uniform$(0,1)$ be independent. Find probability density Function for $Z = X - Y$ and $Z = \min\{X, Y\}$

For $Z = X - Y$

$$F_Z(z) = Pr\{Z \le z\} = Pr\{X - Y \le z\}$$

$$= \begin{cases} \int_0^{1+z} \int_{x-z}^{1} 1 \, dy \, dx & -1 < z < 0 \\ 1 - \int_z^1 \int_0^{x-z} 1 \, dy \, dx & 0 \le z < 1 \end{cases}$$

$$= \begin{cases} \frac{z^2}{2} + z + 1/2 & -1 < z < 0 \\ -\frac{z^2}{2} + z + 1/2 & 0 \le z < 1 \end{cases}$$

By differentiating to get the p.d.f.

$$\boxed{f_Z(z) = \begin{cases} z + 1 & -1 < z < 0 \\ 1 - z & 0 \le z < 1 \end{cases}}$$

For $Z = \min\{X, Y\}$

$$1 - F_Z(z) = P_r\{Z > z\}$$

$$= P_r\{\min\{X, Y\} > z\}$$

$$= P_r\{X > z, Y > z\} \quad X, Y \text{ independent}$$

$$= P_r\{X > z\} P_r\{Y > z\}$$

$$= (1 - z)^2$$

$$F_Z = 1 - (1 - z)^2$$

$$= \begin{cases} 1 - (1 - z)^2 & \text{for } 0 \le z \le 1 \\ 0 & \text{else.} \end{cases}$$

differentiating to get p.d.f.

$$\Rightarrow f_Z(z) = \begin{cases} 2(1 - z) & 0 \le z \le 1 \\ 0 & \text{else.} \end{cases}$$

Problem 1D.

Show that $\mathrm{Var}(Y) = \mathbb{E}\,\mathrm{Var}(Y|X) + \mathrm{Var}\,\mathbb{E}(Y|X)$

$$\mathbb{E}\,\mathrm{Var}(Y|X) + \mathrm{Var}\,\mathbb{E}(Y|X) = \mathbb{E}\big(\mathbb{E}(Y^2|X)\big) - \cancel{\mathbb{E}\big((\mathbb{E}(Y|X))^2\big)} - \big(\mathbb{E}(\mathbb{E}(Y|X))\big)^2 + \cancel{\mathbb{E}\big((\mathbb{E}(Y|X))^2\big)}$$

$$= \mathbb{E}(Y^2) - (\mathbb{E}Y)^2 = \mathrm{Var}\,Y.$$

Problem 2a.

show $\hat{y} = Hy = X\hat{\beta}$ are the least squares estimates

$$y = Xb + e \qquad e = y - Xb.$$

$$S(b) = \Sigma e_i^2 = (y - Xb)'(y - Xb) = y'y - y'Xb - b'X'y + b'X'Xb.$$

take derivative to find minimum. $\frac{\partial S}{\partial b} = -2X'y + 2X'Xb$ <u>minimize S.</u>

$$0 = -\cancel{2}X'y + \cancel{2}X'Xb$$

$$X'Xb = X'y.$$

$$\underline{b = (X'X)^{-1}X'y}$$

Problem 2b.

Show that $HX = X$.

$$H = X(X^T X)^{-1} X^T$$

$$HX = X\cancel{(X^T X)^{-1}}\cancel{(X^T X)} = X.$$

Problem 2c.

Show that H is symmetric : $H = H^T$

$$X(X^TX)^{-1}X^T = \text{transpose}[\,X(X^TX)^{-1}X^T\,]$$

$(X^TX)^{-1}$ is symmetric.

$$= X\left(\left(X^TX\right)^{-1}\right)'X^T$$

since X^TX^{-1} is symmetric.

$$= \boxed{X(X^TX)^{-1}X^T}$$ same as H.

Problem 2D.

show : $H^2 = H$.

$$H \cdot H = X(X'X)^{-1}X' \cdot X(X'X)^{-1}X'$$

$$= X\,\cancel{(X'X)^{-1}}\,\cancel{(X'X)}(X'X)^{-1}X'$$

$$= X(X'X)^{-1}X' = H.$$

Problem (2e)

Show that $\hat{y} = Hy$ is the projection of y onto column space $\mathcal{L}$.

y are vector of dimensions d.

$$X\beta = \sum_{i=1}^{d} \beta_i x_i$$

least squares estimator w/ μ unknown parameter.

$$S(\mu) = \| y - \mu \|^2 = (y-\mu)'(y-\mu)$$

μ spans over column space $\mathcal{L}$

if $X'X$ invertible & assuming colums of X are linearly independent.

then, $\hat{\mu} = X\hat{\beta} = X(X'X)^{-1}X'y$

$$= Hy.$$

Problem (2f)

$X \in \mathbb{R}^{n \times d}$.

then $\operatorname{rank}(X) = d$.

$$tr(H) = tr(X(X^TX)^{-1}X^T)$$

Because of commutativity of trace

$$= tr(X^TX(X^TX)^{-1}) = tr(I_d) = d.$$