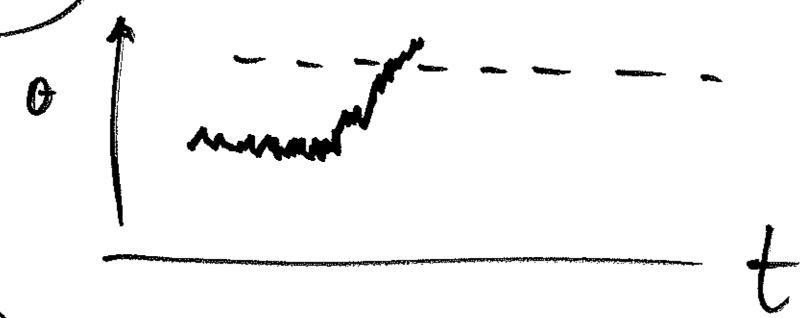
AMS 131
11 Jul 18

1

this multivariate time: distributions (9.51)

next time: ↓



more realistic

goal $(B_i \mid \theta_i) \overset{I}{\sim} \text{Bernoulli}_i(\theta_i)$

$(i = 1, \ldots, n)$

but this is unattainable:

n data points, n unknowns (too many unknowns)

$$\prod_{i=1}^{n} \theta^{b_i} (1-\theta)^{1-b_i} =$$

$$\left[\theta^{b_1}(1-\theta)^{1-b_1}\right]\left[\theta^{b_2}(1-\theta)^{1-b_2}\right] \cdots \left[\theta^{b_n}(1-\theta)^{1-b_n}\right]$$

$$= \theta^{b_1 + \cdots + b_n} (1-\theta)^{n-(b_1 + \cdots + b_n)}$$

$$= \theta^{s}(1-\theta)^{n-s}$$

pts. in T w/ good outcome $\rightarrow s = \sum_{i=1}^{n} b_i$

Y continuous RV $(n=1)$ with CDF

$F_Y(y)$ and PDF $f_Y(y)$, then

$$f_Y(y) = \frac{d}{dy} F_Y(y)$$

$\underset{\sim}{Y} = (Y_1, Y_2)$ continuous random vector

$(n=2)$ with CDF $F_{\underset{\sim}{Y}}(\underset{\sim}{y})$ and

PDF $f_{\underset{\sim}{Y}}(\underset{\sim}{y})$, then $\qquad F_{Y_1,Y_2}(y_1, y_2)$

$$f_{\underset{\sim}{Y}}(\underset{\sim}{y}) = \frac{d}{dy_1}\frac{d}{dy_2} F_{Y_1,Y_2}(y_1, y_2)$$

$$= \frac{d}{dy_2}\frac{d}{dy_1} F_{Y_1,Y_2}(y_1, y_2)$$

(10.43)

$Y_i = \begin{cases} 1 & \text{with prob } \theta \\ 0 & 1-\theta \end{cases}$

↑ machine i^{th} outcome

$n = 7$

θ unknown

$y = (0, 1, 1, 1, 0, 1, 1)$

$\frac{5}{7}$

θ known $= 0.75$ ←

$y = (0, 1, 1, 1, 0, 1, 1)$

the Y_i are conditionally independent given θ

data

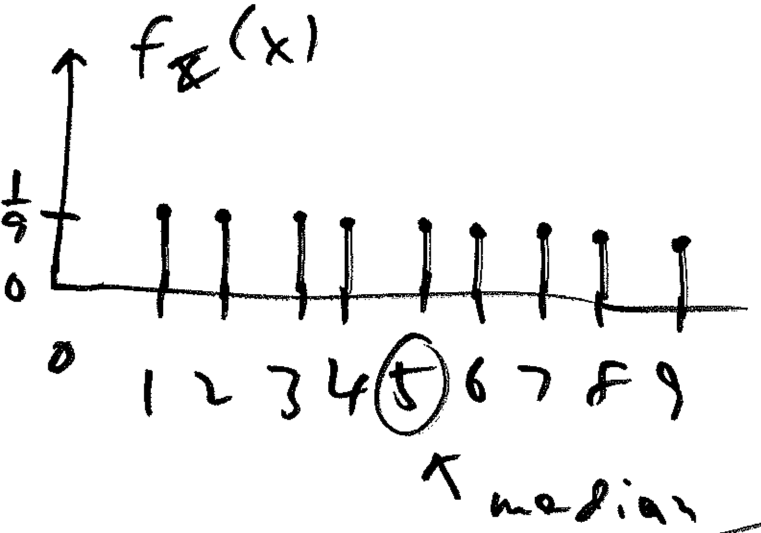
$\underset{\sim}{Y} = (Y_1, \dots, Y_n)$

$(Y_i \mid \theta)$ are cond. indep.

unknown

then $f_{\underset{\sim}{Y}}(y \mid \theta) = \prod_{i=1}^{n} f_{Y_i}(y_i \mid \theta)$

crucial ingredient in statistical inference

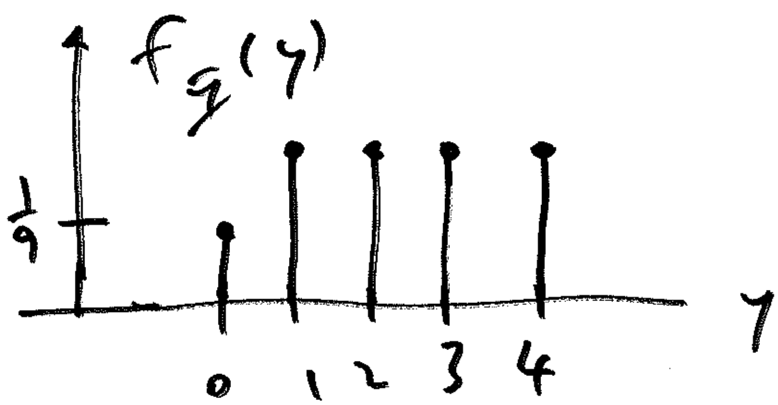


PMF of X

$Y = h(X)$

$= |X - 5|$

possible values y for Y	$X = x$ such that $Y = \lvert X - 5 \rvert = y$	prob. $P(Y = y)$
0	5	$\frac{1}{9}$
1	4 or 6	$\frac{2}{9}$
⋮	⋮	⋮



PMF of Y