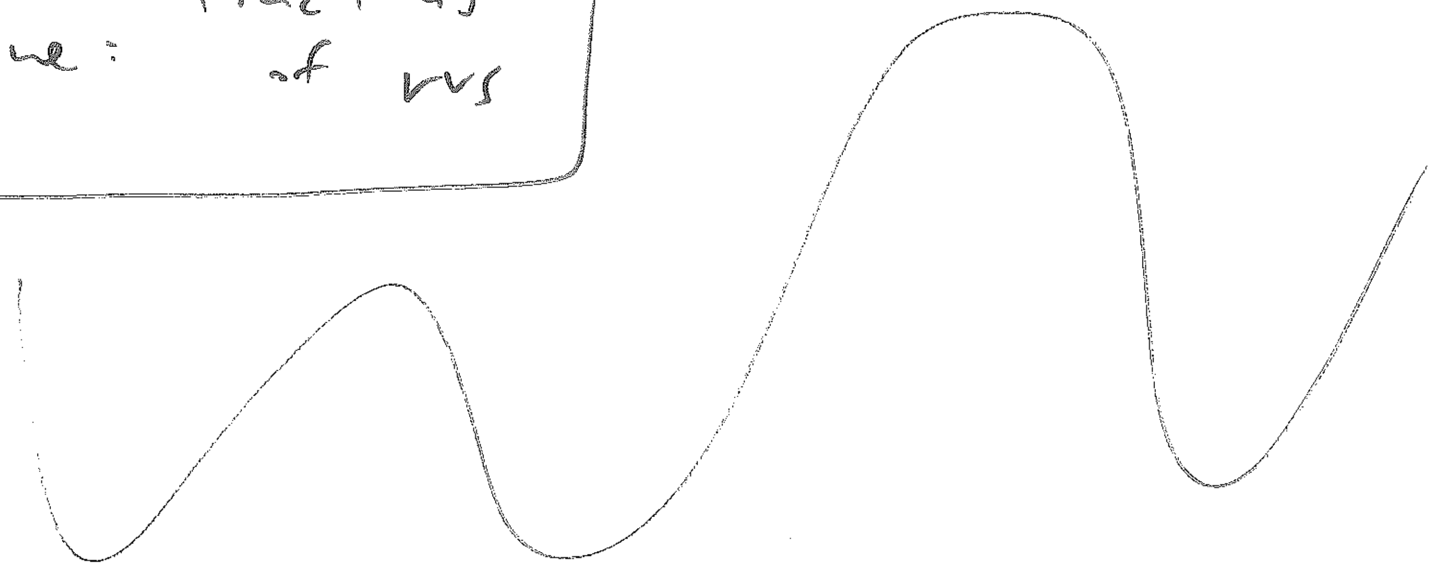
AMS 13
14 Aug 17

1

this time: multivariate distributions

next time: functions of rvs

read: finish DS ch. 3, start ch. 4



Case study: gender & marijuana legalization preference (MLP) at UCLA

raw data

(Y) MLP	gender (X)
Y	F
Y	M
N	M
⋮	⋮

106

sort

Y ... Y	F ... F	29
N ... N	F ... F	20
Y ... Y	M ... M	52
N ... N	M ... M	5

cross-tabulation of gender & MLP: 2 x 2 contingency table

gender (X)	MLP (Y): (1) Y	(0) N	
(1) F	29	20	49
(0) M	52	5	57
	81	25	106

"pick one of these 106 people at random" ②

$$\underline{X} = \begin{cases} 1 & \text{if } F \\ 0 & M \end{cases}$$

$$\underline{Y} = \begin{cases} 1 & \text{if } Y \\ 0 & N \end{cases}$$

joint (mass) PF (prob. mass fn) pf

$(\underline{Y}, \underline{X})$ bivariate discrete rv

PF $f_{\underline{Y},\underline{X}}(y, x) = ⊛$

and

⊛ $P(\underline{Y} = 1, \underline{X} = 1) = \frac{29}{106} \doteq 27\%$

$P(\underline{Y} = 1, \underline{X} = 0) = \frac{52}{106} \doteq 49\%$

$P(\underline{Y} = 0, \underline{X} = 1) = \frac{20}{106} \doteq 19\%$

$P(\underline{Y} = 0, \underline{X} = 0) = \frac{5}{106} \doteq 5\%$

100%

marginal (M) PF for $\underline{X}$:

$$f_{\underline{X}}(x) = \sum_{\text{all } y} P(\underline{Y} = y, \underline{X} = x)$$

$$f_{\underline{X}}(1) = P(\underline{X} = 1) = \frac{49}{106} = \overbrace{P(\underline{Y} = 0, \underline{X} = 1)}^{20/106} + \overbrace{P(\underline{Y} = 1, \underline{X} = 1)}^{29/106}$$

$$P(Y = 1) = \frac{81}{106} \doteq 76\%$$

$$P(Y = 1 \mid X = 0) = \frac{52}{57} \doteq 91\%$$

$$P(Y = 1 \mid X = 1) = \frac{29}{49} \doteq 59\%$$

Because 91%, 59% differ from 76% by so much, gender & MCP are strongly dependent in this dataset.

Clinical trial example p. 128 of extra notes

ideally $0 < \theta_i < 1$

$$P(B_i = b_i) = \begin{cases} 1 & \text{with prob } \theta_i \\ 0 & \text{——— } (1 - \theta_i) \end{cases}$$

for scientific realism, want everybody to have her/his own θ_i but can't learn $(\theta_1, \dots, \theta_n)$ from a dataset of the form $(B_1, \dots, B_n)$

(continuation of p. 133 extra notes)

(nut/bolt case study continued)

$$p(\theta \mid N) = \frac{p(\theta)\, p(N \mid \theta)}{p(N)}$$

posterior dist. for θ given N

prior distribution for θ

sampling distribution

normalizing constant

so $p(\theta \mid N) \propto p(\theta)\, p(N \mid \theta)$

is proportional to

$$\binom{m}{n} \theta^{n} (1-\theta)^{m-n}$$

Bayes (1760): if choose prior of form $\theta^{\alpha-1}(1-\theta)^{\beta-1}$, post. has same form as prior $\theta^{\alpha-1}(1-\theta)^{\beta-1}$.

$n = 3$

$m = 114$

$$p(\theta \mid N = n) \propto \boxed{p(\theta)}\; \theta^{n}(1-\theta)^{m-n}$$

(prob. dist for θ) (prob. dist for θ)

Beta distribution $\quad \Theta \sim \text{Beta}(\alpha, \beta)$

$$\rightarrow \; p(\theta) = c\,\theta^{\alpha-1}(1-\theta)^{\beta-1}$$

$\left(f_{\Theta}(\theta) \right)$ ARS 131

if we choose prior to be Beta(α, β)

then posterior =

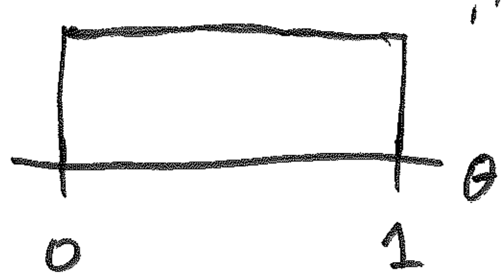
$$p(\theta \mid N=n) \propto \left[\theta^{\alpha-1}(1-\theta)^{\beta-1}\right]\left[\theta^{n}(1-\theta)^{m-n}\right]$$

$$= \theta^{(\alpha+n)-1}(1-\theta)^{(\beta+m-n)-1}$$

$$= \text{Beta}(\overset{1+3}{\alpha+n}, \overset{1+114-3}{\beta+m-n})$$

I don't know much about θ external to data set

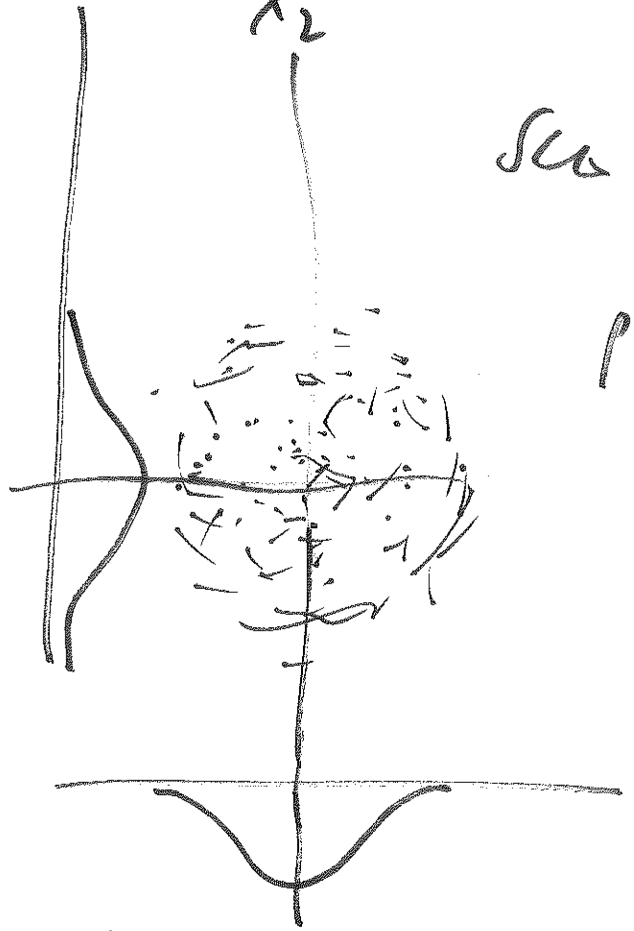
Uniform(0,1)



Uniform(0,1) = Beta(1,1)

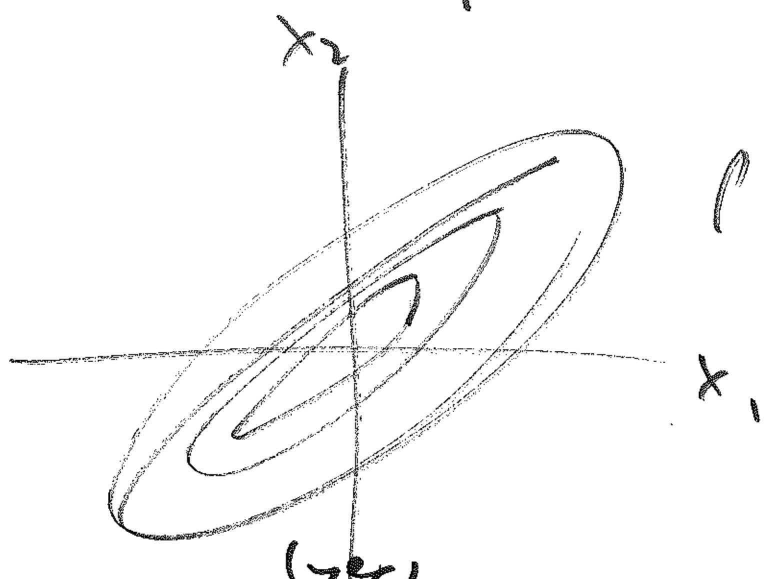
$\begin{pmatrix} \alpha = 1 \\ \beta = 1 \end{pmatrix}$ $\quad p(\theta \mid N = 4) = \text{Beta}(4, 112)$

$$
\begin{array}{c}
\begin{array}{cc} X_1 & X_2 \end{array} \\
\begin{bmatrix}
X_{11} & X_{12} \\
X_{21} & X_{22} \\
\vdots & \\
X_{n1} & X_{n2}
\end{bmatrix}
\end{array}
\quad n
$$



Scatterplot

$\rho = 0$



$\rho = +.9$

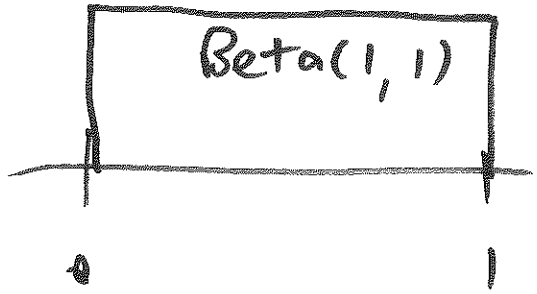
(yes) indep ? ✓ marginal for A marginal for B

$P(A, B) \; \cdot \; \underline{P(A)} \; \underline{P(B)}$

general (no)

$= P(A)\,P(B \mid A) = P(B)\,P(A \mid B)$

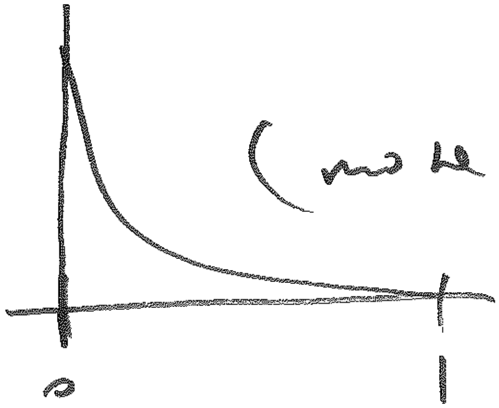
$U(0,1) =$



info. state

$\boxed{\text{I don't know anything about } \theta \text{ except}}$ $0 < \theta < 1$

(neutral prior)



(more realistic prior)

---

unknown

$$\underline{Y} = \begin{pmatrix} Z_1 \\ \underline{Y} \mid Z_1 \end{pmatrix}$$

indirect

$\underline{Y}$ depends on Z_1

$f_{\underline{Y}}(y)$

directly draw random samples from

hierarchical representation of $\underline{Y}$

① sample from $f_{Z_1}(z)$, obtaining z^*

② sample from $f_{\underline{Y} \mid Z_1}(y \mid z^*)$

$$p(\theta \mid \underline{y}) = \frac{p(\theta)\, p(\underline{y} \mid \theta)}{p(\underline{y})}$$

?

truth

hard

$$p(\underline{y}) = \int p(\underline{y}, \theta)\, d\theta$$

$$= \int p(\theta)\, \underline{\underline{p(\underline{y} \mid \theta)}}\, d\theta$$