Lecture 8

marginal, conditional, and joint PDFs

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} \qquad f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

$$\underset{\sim}{Y} = (Y_1, Y_2, Y_3, Y_4) \qquad f_{Y_1}(y_1) = \iiint f_{\underset{\sim}{Y}}(\underset{\sim}{y})\, dy_2\, dy_3\, dy_4$$

integrate with respect to every other one you are not looking for

$$f_{Y_2, Y_3}(y_2, y_3) = \iint f_{\underset{\sim}{Y}}(\underset{\sim}{y})\, dy_1\, dy_4$$

Can get a marginal CDF by sending $\to \infty$

$$F_{Y_1}(y_1) = P(Y_1 \le y_1) = P(Y_1 \le y_1, Y_2 < \infty, \ldots Y_n < \infty)$$

$$= \lim_{y_2 \to \infty \ldots y_n \to \infty} F_{\underset{\sim}{Y}}(\underset{\sim}{y})$$

n R.V.s are independent if for any non weird sets $A_1, \ldots, A_n$ of real numbers

$$P(Y_1 \in A_1, \ldots, Y_n \in A_n) = \prod_{i=1}^{n} P(Y_i \in A_i)$$

(joint = product of the marginals)

- $F_{\underset{\sim}{Y}}(\underset{\sim}{y}) = \prod_{i=1}^{n} F_{Y_i}(y_i)$  - $f_{\underset{\sim}{Y}}(\underset{\sim}{y}) = \prod_{i=1}^{n} f_{Y_i}(y_i)$

$$\text{pop.}\ \begin{bmatrix} \ \\ \ \\ \ \end{bmatrix} \xrightarrow[\text{IID}]{} \begin{matrix} Y_1 \\ Y_2 \\ \cdots \\ Y_n \end{matrix} \overset{\text{sample}}{\begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix}} \Big\} n$$

$\underset{\sim}{Y} = (Y_1, \ldots, Y_n)$ is a random sample from population.

Starting with a univariate PMF/PDF $f_{Y_i}(y_i)$ of size n, R.V.s $(Y_1, \ldots, Y_n)$ form a random sample from $f_{Y_i}(y_i)$ if all of them have marginal PMF/PDF $f_{Y_i}$ $\leftrightarrow$ if $Y_i$ are an IID sample from $f_{Y_i}$

AMS 131

# Lecture 8 (cont.)

ex: deer and chronic wasting disease

all deer $\qquad$ observed deer $\qquad$ 1 = Yes, 0 = No

$$N \approx 800 \begin{bmatrix} 1\text{'s} \\ \text{and} \\ 0\text{'s} \end{bmatrix} \xrightarrow{\text{IID}} \left.\begin{bmatrix} 1\text{'s} \\ \text{and} \\ 0\text{'s} \end{bmatrix}\right\} n = \begin{bmatrix} Y_1 \\ \cdots \\ Y_n \end{bmatrix}$$

disease? $\qquad$ mean $\bar{y} = \hat{\theta}$ (estimate of)

mean $\theta = ?$ (unknown) $\qquad$ Sample mean represents whole $\theta$

IID good sampling technique: $\hat{\theta} \leftrightarrow \theta$

Short hand for IID diagram: $(Y_i \mid \theta) \overset{\text{IID}}{\sim} \text{Bernoulli}(\theta)$ $\quad (i = 1, \ldots, n)$

$\underset{\sim}{X} = (X_1, \ldots, X_n)$ random vector partitioned into 2 subvectors $\underset{\sim}{X} = (\underset{\sim}{Y}, \underset{\sim}{Z})$

$\underset{\sim}{Y} = (Y_1, \ldots, Y_k)$ $\quad 1 \le k \le n-1$

$\underset{\sim}{Z} = (Z_1, \ldots, Z_{n-k})$ remaining elements

then for every point $\underset{\sim}{z}$ (lowercase) for which $f_{\underset{\sim}{Z}}(\underset{\sim}{z}) > 0$ the conditional distribution of $\underset{\sim}{Y}$ given $\underset{\sim}{Z}$ is:

$$f_{\underset{\sim}{Y} \mid \underset{\sim}{Z}}(\underset{\sim}{y} \mid \underset{\sim}{z}) = \frac{f_{\underset{\sim}{Y}, \underset{\sim}{Z}}(\underset{\sim}{y}, \underset{\sim}{z})}{f_{\underset{\sim}{Z}}(\underset{\sim}{z})}, \quad \underset{\sim}{y} \in \mathbb{R}^k$$

from which:

$$f_{\underset{\sim}{Y}, \underset{\sim}{Z}}(\underset{\sim}{y}, \underset{\sim}{z}) = f_{\underset{\sim}{Z}}(\underset{\sim}{z}) \cdot f_{\underset{\sim}{Y} \mid \underset{\sim}{Z}}(\underset{\sim}{y} \mid \underset{\sim}{z})$$

## Multivariate Law of Total Probability


$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(B_i)\, P(A \mid B_i)$$

$\updownarrow$ for R.V.s

$$\underbrace{f_{\underset{\sim}{Y}}(\underset{\sim}{y})}_{\text{"A"}} = \int \cdots \int_{\mathbb{R}^{n-k}} \underbrace{f_{\underset{\sim}{Z}}(\underset{\sim}{z})}_{\text{"Bi"}} \cdot \underbrace{f_{\underset{\sim}{Y} \mid \underset{\sim}{Z}}(\underset{\sim}{y} \mid \underset{\sim}{z})}_{\text{"A|Bi"}}\, d\underset{\sim}{z}$$

mixture: weighted average with $\underset{\sim}{Z}$

Lecture 8 (cont.) AMS 131

$$f_{\underset{\sim}{Z}|\underset{\sim}{Y}}(\underset{\sim}{z}|\underset{\sim}{y}) = \frac{f_{\underset{\sim}{Z}}(\underset{\sim}{z}) \cdot f_{\underset{\sim}{Y}|\underset{\sim}{Z}}(\underset{\sim}{y}|\underset{\sim}{z})}{f_{\underset{\sim}{Y}}(\underset{\sim}{y})}$$

posterior information = $\frac{\text{prior info} \cdot \text{likelihood}}{\text{normalizing constant}}$

$\underset{\sim}{Z}$ a random vector with multivariate distribution $f_{\underset{\sim}{Z}}(\underset{\sim}{z})$
then R.V.s $X_1, \ldots, X_n$ are conditionally independent given $\underset{\sim}{Z}$
if for all $\underset{\sim}{z}$ with $f_{\underset{\sim}{Z}}(\underset{\sim}{z}) > 0$:

$$f_{\underset{\sim}{X}|\underset{\sim}{Z}}(\underset{\sim}{x}, \underset{\sim}{z}) = \prod_{i=1}^{n} f_{X_i|\underset{\sim}{Z}}(x_i|\underset{\sim}{z})$$

machine: IID coin tosses  P(heads) = $\theta$
if $\theta$ is unknown: the results $Y_1, Y_2 \ldots$ are dependent, because there is useful information for predicting future subsets (7 heads, $\theta = \frac{7}{10}$)
if $\theta$ is known: the knowledge of the truth overrides data, they became conditionally independent given $\theta$
ex: the clinical trial and nuts/bolts, we model data $Y_i$ as $(Y_i|\theta) \overset{IID}{\sim}$ Bernoulli($\theta$) conditionally.

Functions of a random variable

1. discrete: $X$ (univariate) discrete R.V  PMF = $f_X(x)$
   $Y = h(X)$ - some function $h$ defined on {possible values of $X$}

$$f_Y(y) = P(Y=y) = P[h(X)=y]$$
$$= \sum_{x:h(x)=y} f_X(x)$$ add up all PMF along all x's where $h(x) = y$

$X \sim$ Uniform$\{1, 2, \ldots, 9\}$  median: 5
$Y = |X-5| = h(X)$ keeps track of distance to median

| y | $X: Y=y$ | $P(Y=y)$ |
|---|---|---|
| 0 | 5 | $\frac{1}{9}$ |
| 1 | 4 or 6 | $\frac{2}{9}$ |
| 2 | 3 or 7 | $\frac{2}{9}$ |
| 3 | 2 or 8 | $\frac{2}{9}$ |
| 4 | 1 or 9 | $\frac{2}{9}$ |

AMS 131 Lecture 8 (cont.)

$X$ is now continuous R.V. PDF $= f_X(x)$; $Y = h(X)$

The CDF of $F_Y(y)$: $F_Y(y) = P(Y \le y) = P[h(X) \le y]$

$$= \int_{\{x:\, h(x) \le y\}} f_X(x)\,dx$$ if $Y$ also continuous $f_Y(y) = \frac{d}{dy} F_Y(y)$ = PDF of $Y$

ex: $X$ = rate customers are served in a line at a bank
- continuous, $X > 0$, CDF $F_X$

The average wait time: $Y = \frac{1}{X} = h(X)$ you can get the PDF of $Y$:

1. work out the CDF of $Y$
2. differentiate with respect to $y$

1. for $y > 0$ $F_Y(y) = P(Y \le y) = P[h(X) \le y]$

$$= P\left[\frac{1}{X} \le y\right] = P\left[X \ge \frac{1}{y}\right]$$

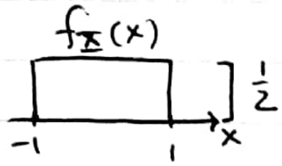
$$F_Y(y) = 1 - P(X < \tfrac{1}{y}) = 1 - P(X \le \tfrac{1}{y}) = 1 - F_X(\tfrac{1}{y})$$

2. $f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy}\left[1 - F_X(\tfrac{1}{y})\right]$ chain rule

$$= -f_X(\tfrac{1}{y})(-y^{-2}) = \frac{f_X(\frac{1}{y})}{y^2}$$

ex: $X \sim$ Uniform $[-1, 1]$ $Y = X^2$ find PDF $Y$

Support of $Y$ $[0, 1]$



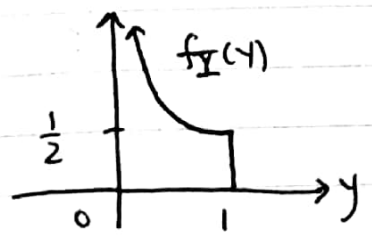

1. $F_Y(y) = P(Y \le y) = P(X^2 \le y)$ for $0 < y < 1$

$$= P(-\sqrt{y} \le X \le \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)\,dx = \frac{1}{2}x\Big|_{-\sqrt{y}}^{\sqrt{y}}$$

$$f_X(x) = \begin{cases} \frac{1}{2} & -1 \le x \le 1 \\ 0 & \text{else} \end{cases}$$

$F_Y(y) = \sqrt{y}$

2. $f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{1}{2\sqrt{y}}$ $0 < y < 1$




the density is unbounded at $y = 0$

Lecture 8 (cont.) AMS 131

$X$ is continuous R.V. with PDF $f_X(x)$

$Y = aX + b$ linear transformation $a \neq 0$

$\rightarrow f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$

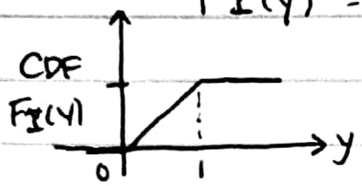
CDF $F_X(x)$

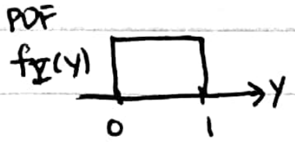
input R.V!

what is $Y = F_X(X)$?

$F_Y(y) = P(Y \leq y) = P[F_X(X) \leq y] = P[X \leq F_X^{-1}(y)]$

$F_Y(y) = F_X[F_X^{-1}(y)] = y$ for $0 < y < 1$




if you feed $F_X(X)$, the new variable is uniform

$Y \sim \text{Uniform}(0,1)$

$X \sim F_X(x)$ $Y = F_X(X) \sim U(0,1)$ $\rightarrow X = F_X^{-1}[U(0,1)]$

Probability Integral Transform

$X$ continuous with CDF $F_X$

$Y = F_X(X) \rightarrow Y \sim \text{Uniform}(0,1)$ or $[0,1]$

The converse is also true: $Y \sim \text{Uniform}[0,1]$

$F_X$ is a continuous CDF with quantile function

$F_X^{-1} \rightarrow X = F_X^{-1}(Y) \sim F_X$

- useful for computing pseudo-random numbers, since pseudo-uniform $(0,1)$ are easy to generate = $F_X^{-1}$ easy/fast compute.

$U_1, \dots, U_n \overset{IID}{\sim} \text{Uniform}(0,1)$

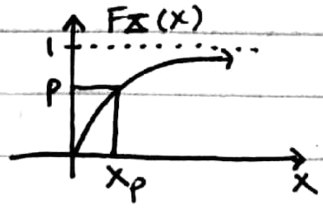
$F_X^{-1}(U_1), \dots, F_X^{-1}(U_n) \overset{IID}{\sim} F_X$

family of exponentials where $\lambda > 0$

If $X \sim \text{Exponential}(\lambda)$ $\lambda > 0$

$$\begin{cases} \lambda e^{-\lambda x} & \text{for } x > 0 \\ 0 & \text{else} \end{cases} \quad \text{PDF}$$

for $x > 0$ $F_X(x) = 1 - e^{-\lambda x}$




$$\begin{cases} 0 & x \leq 0 \\ 1 - e^{-\lambda x} & x > 0 \end{cases} \quad \text{CDF}$$

$F_X(x_p) = p$ so,

$p = 1 - e^{-\lambda x_p}$ $e^{-\lambda x_p} = 1 - p$ $-\lambda x_p = \ln(1-p)$ $x_p = \frac{\log(1-p)}{-\lambda} = F_X^{-1}(p)$

$(X_i^* | \lambda) \overset{IID}{\sim} \text{Exponential}(\lambda)$ $U_i^* \sim \text{Uniform}(0,1)$

$X_i^* = -\frac{\log(1-U_i^*)}{\lambda} \rightarrow -\frac{1}{\lambda}\log(U_i)$

AMS 131

Lecture 8 (cont.)

Some situations are too complicated to characterize mathematically in closed form, so we can conduct a simulation study, driven by pseudo-random

- simulation: you don't have closed form (mathematical entity vs. numbers)

Nice $h(X)$ functions are differentiable and one-to-one (invertible)

If $h(x)$ is differentiable and 1-1 for $x$ in the open interval $(a,b)$, then $h$ is either montonically increasing or decreasing

$h$ is also continuous so it transforms $(a,b)$ to $(\alpha,\beta)$ - the image of $(a,b)$ under $h$

$y = h(x) \leftrightarrow x = h^{-1}(y)$ invertibility

$X$ continuous R.V. with PDF $f_X(x)$ and for which $P(a < X < b) = 1$ $\quad Y = h(X)$ with $h$ differentiable and 1-1 for $a < x < b$ $\quad (\alpha,\beta)$ image of $(a,b)$

$h^{-1}(y)$ inverse of $h(x)$ for $\alpha < y < \beta$

$$\text{PDF } f_Y(y) = \begin{cases} f_X[h^{-1}(y)] \left| \dfrac{dh^{-1}(y)}{dy} \right| & \text{for } \alpha < y < \beta \\ 0 & \text{else} \end{cases}$$

$f_Y(y)|dy| = f_X(x)|dx|$ convert from $x \rightarrow y$ with no loss

$Y = \frac{1}{X} = h(X)$ $\quad y = h(x) = \frac{1}{x}$, $x = h^{-1}(y) = \frac{1}{y}$

$\frac{d}{dy}\frac{1}{y} = -\frac{1}{y^2}$ $\quad f_Y(y) = \frac{f_X(\frac{1}{y})}{y^2}$ as before

Lecture 8 (cont.) AMS 131

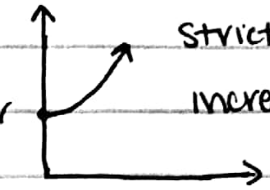
At time = 0, a population of organisms is introduced into a tank with nutrients – 'v' organisms

$X$ = rate of growth (unknown) $\quad Y = ve^{Xt}$ (exponential)

$$f_X(x) = \begin{cases} 3(1-x^2) & 0 < x < 1 \\ 0 & \text{else} \end{cases}$$

$y = h(x) = ve^{xt}$ $\quad x = 0 \rightarrow y = v$

$x = 1 \rightarrow y = ve^{xt}$ (image)




strictly increasing

$$\frac{y}{v} = e^{xt} \rightarrow \log\left(\frac{y}{v}\right) = xt \rightarrow x = h^{-1}(y) = \frac{1}{t}\log\left(\frac{y}{v}\right)$$

$$\frac{d}{dy}\frac{1}{t}\log\left(\frac{y}{v}\right) = \frac{1}{t}\left(\frac{y}{v}\right)^{-1}\cdot\frac{1}{v} = \frac{1}{ty} \qquad f_Y(y) = \begin{cases} \frac{3}{ty}\left[1 - \frac{1}{t}\log\left(\frac{y}{v}\right)\right]^2 \\ v < y < ve^t & 0 \text{ else} \end{cases}$$

this resembles the form $\quad f_Y(y) = \begin{cases} f_X(\text{"}x\text{"}) \left|\frac{dx}{dy}\right| & \alpha < y < \beta \\ 0 & \text{else} \end{cases}$

Functions with two or more R.V.s $\quad$ transform $X$ to $Y$

1. Discrete: $n$ R.V.s $X_1, \dots, X_n$

discrete joint distribution with joint PDF $f_{\underset{\sim}{X}}(\underset{\sim}{x})$

define: $\begin{cases} Y_1 = h_1(X_1, \dots, X_n) \\ \dots \qquad \dots \\ Y_m = h_m(X_1, \dots, X_n) \end{cases}$ $\quad m \geq 1 \quad (h_j : \mathbb{R}^n \rightarrow \mathbb{R})$ real-valued

$\underset{\sim}{y} = (y_1, \dots, y_m)$ of $\underset{\sim}{Y} = (Y_1, \dots, Y_m)$ $\quad A = (x_1, \dots, x_n)$ such that

$$\begin{cases} y_1 = h_1(x_1, \dots, x_n) \\ \dots \qquad \dots \\ y_m = h_m(x_1, \dots, x_n) \end{cases} \rightarrow \text{joint PMF } f_{\underset{\sim}{Y}}(\underset{\sim}{y}) = \sum_{(x_1, \dots, x_n) \in A} f_{\underset{\sim}{X}}(\underset{\sim}{x})$$

2. Continuous: $n$ R.V.s $X_1, \dots, X_n$ cont. joint dist. PDF $f_{\underset{\sim}{X}}(\underset{\sim}{x})$

$Y = h(\underset{\sim}{X})$ $\quad$ For each real $y$, define: $A_y = \{\underset{\sim}{x} : h(\underset{\sim}{x}) = y\}$

univariate (real) $m = 1$ $\quad f_Y(y) = \int_{A_y} \cdots \int f_{\underset{\sim}{X}}(\underset{\sim}{x}) \, d\underset{\sim}{x}$

ex: $(X_1, X_2)$ joint continuous PDF $f_{X_1, X_2}(x_1, x_2)$

$Y = a_1X_1 + a_2X_2 + b$ with $a_1 \neq 0 \rightarrow Y$ continuous PDF

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1, X_2}\left(\frac{y - b - a_2x_2}{a_1}, x_2\right)\frac{dx_2}{|a_1|}$$

AMS 131

Lecture 8 (cont.)

Special case: The simplest thing you can do with two or more R.V.s is to add them.

sample mean: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ plays key role in statistics

$Y = X_1 + X_2 \qquad (a_1, a_2, b) = (1, 1, 0)$

The distribution of $Y$ is the convolution of $X_1$ and $X_2$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1}(y-z) f_{X_2}(z)\,dz = \int_{-\infty}^{\infty} f_{X_1}(z) f_{X_2}(y-z)\,dz$$

$X_i \overset{IID}{\sim}$ CDF $F_{X_i}$, PDF $f_{X_i}$ $(i = 1, 2, 3 \ldots n)$ continuous

$Y_{(1)} \triangleq \min(X_1, \ldots, X_n)$
$Y_{(n)} \triangleq \max(X_1, \ldots, X_n)$
} order statistics of $(X_1, \ldots, X_n)$

$F_{Y_{(n)}}(t) = P(Y_{(n)} \leq t)$ CDF of maximum point

$\leftrightarrow$ iff $P(X_1 \leq t, X_2 \leq t, \ldots, X_n \leq t)$

$= P(X_1 \leq t) \ldots P(X_n \leq t)$ independent IID

$F_{Y_{(n)}}(t) = [F_{X_i}(t)]^n$ IID

So $Y_{(n)}$ has a PDF of: $f_{Y_{(n)}}(t) = \frac{d}{dt}\left[F_{X_i}(t)\right]^n = n\left[F_{X_i}(t)\right]^{n-1} f_{X_i}(t)$

$F_{Y_{(1)}}(t) = P(Y_{(1)} \leq t) = 1 - P(Y_{(1)} > t)$ IID

$\leftrightarrow$ iff $1 - P(X_1 > t, \ldots, X_n > t) = 1 - P(X_1 > t) \ldots P(X_n > t)$

$= 1 - [1 - F_{X_i}(t)]^n$ so $Y_{(1)}$ PDF is:

$$f_{Y_{(1)}}(t) = \frac{d}{dt} F_{Y_{(1)}}(t) = n[1 - F_{X_i}(t)]^{n-1} f_{X_i}(t)$$

Lecture 8 (cont.) AMS 131

Multivariate transformations

$X_1, \dots, X_n$ cont. joint dist. PDF $f_{\underset{\sim}{X}}(\underset{\sim}{x})$

Suppose there is a subset $S$ (the support of $(X_1, \dots, X_n)$)
with $P[(X_1, \dots, X_n) \in S] = 1$ define new R.V.s

$Y_1 = h_1(X_1, \dots, X_n)$ assume the $n$ functions
$\dots \quad \dots$ $h_1, \dots, h_n$ are differentiable
$Y_n = h_n(X_1, \dots, X_n)$ and 1-1

transformation of $S$ onto some subset $T$ of $\mathbb{R}^n$
(support of the image of $h_1, \dots, h_n$)

$$\left.\begin{aligned} x_1 &= h_1^{-1}(y_1, \dots, y_n) \\ &\dots \\ x_n &= h_n^{-1}(y_1, \dots, y_n) \end{aligned}\right\} \text{inverse transformation}$$

$$f_{\underset{\sim}{Y}}(\underset{\sim}{y}) = \begin{cases} f_{\underset{\sim}{X}}[h_1^{-1}(\underset{\sim}{y}), \dots, h_n^{-1}(\underset{\sim}{y})]\,|J| & (y_1, \dots, y_n) \in T \\ 0 & \text{else} \end{cases}$$

$J$ is the anolog of the derivative when $X \to Y$ $\quad \frac{dh^{-1}(y)}{dy}$

$J$ is the determinant of the matrix: $|J|$ abs. value

$$\begin{bmatrix} \frac{dh_1^{-1}}{dy_1} & \dots & \frac{dh_1^{-1}}{dy_n} \\ \dots & & \dots \\ \frac{dh_n^{-1}}{dy_1} & \dots & \frac{dh_n^{-1}}{dy_n} \end{bmatrix}$$

Called the Jacobian of the transformation from $\underset{\sim}{X}$ to $\underset{\sim}{Y}$.
It acts like a generalization of the derivative of the inverse

$(X_1, X_2)$ joint, continuous

$$\text{PDF: } f_{X_1 X_2}(x_1, x_2) \begin{cases} 4x_1x_2 & 0 < x_1 < 1,\ 0 < x_2 < 1 \\ 0 & \text{else} \end{cases}$$

$(2x_1)(2x_2) \leftarrow$ independent

check: (joint PDF)

$$\int_0^1 \int_0^1 4x_1x_2\, dx_1\, dx_2 = \int_0^1 4x_2 \left(\int_0^1 x_1\, dx_1\right) dx_2$$

$$= 4\int_0^1 x_2 \left[\tfrac{1}{2}x_1^2 \Big|_0^1\right] dx_2 = 2\int_0^1 x_2\, dx_2 = 2\left[\tfrac{1}{2}x_2^2\Big|_0^1\right] = 1$$

$$(Y_1, Y_2) \triangleq \left(\frac{X_1}{X_2},\ X_1 \cdot X_2\right)$$

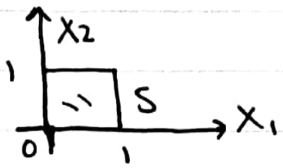
$y_1 = h_1(x_1, x_2) = \frac{x_1}{x_2}$

$y_2 = h_2(x_1, x_2) = x_1 \cdot x_2$

AMS 131 Lecture 8 (cont.)

inverse transform: solve for $(X_1, X_2)$ $\begin{cases} \frac{X_1}{X_2} = Y_1 \\ X_1 X_2 = Y_2 \end{cases}$ $X_1 = h_1^{-1}(Y_1, Y_2) = \sqrt{Y_1 Y_2}$, $X_2 = h_2^{-1}(Y_1, Y_2) = \sqrt{\frac{Y_2}{Y_1}}$

image: how does $0 < X_1 < 1$, $0 < X_2 < 1$ transform?

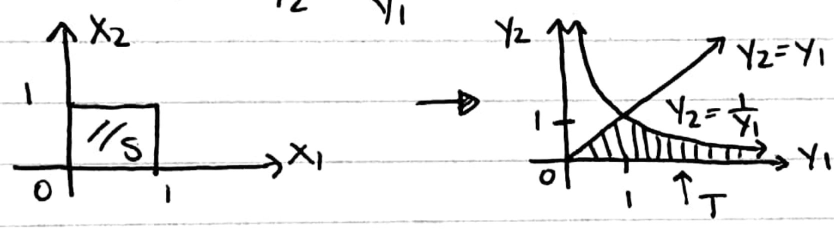



$\begin{cases} X_1 > 0, & X_1 < 1 \\ X_2 > 0, & X_2 < 1 \end{cases}$ so, solve for image

$\sqrt{Y_1 Y_2} > 0$ (a) $\quad \sqrt{Y_1 Y_2} < 1$ (b) $\quad \sqrt{\frac{Y_2}{Y_1}} > 0$ (c) $\quad \sqrt{\frac{Y_2}{Y_1}} < 1$ (d)

(a) $\rightarrow \begin{pmatrix} Y_1 > 0 \\ Y_2 > 0 \end{pmatrix}$ $\begin{pmatrix} Y_1 < 0 \\ Y_2 < 0 \end{pmatrix}$ but $Y_1 = \frac{X_1}{X_2} > 0$ so it must be: $\begin{pmatrix} Y_1 > 0 \\ Y_2 > 0 \end{pmatrix}$ (c) says the same.

(b) says $Y_2 < \frac{1}{Y_1}$ (d) says $Y_2 < Y_1$




(a), (c) means both positive (I).

$h_1^{-1}(Y_1, Y_2) = \sqrt{Y_1 Y_2}$ $\quad \frac{d}{dY_1} h_1^{-1} = \frac{1}{2}\sqrt{\frac{Y_2}{Y_1}}$ $\quad \frac{d}{dY_2} h_2^{-1} = \frac{1}{2}\sqrt{\frac{1}{Y_1 Y_2}}$

$h_2^{-1}(Y_1, Y_2) = \sqrt{\frac{Y_2}{Y_1}}$ $\quad \frac{d}{dY_2} h_1^{-1} = \frac{1}{2}\sqrt{\frac{Y_1}{Y_2}}$ $\quad \frac{d}{dY_1} h_2^{-1} = -\frac{1}{2}\sqrt{\frac{Y_2}{Y_1^3}}$

$$J = \det \begin{bmatrix} \frac{1}{2}\sqrt{\frac{Y_2}{Y_1}} & \frac{1}{2}\sqrt{\frac{Y_1}{Y_2}} \\ -\frac{1}{2}\sqrt{\frac{Y_2}{Y_1^3}} & \frac{1}{2}\sqrt{\frac{1}{Y_1 Y_2}} \end{bmatrix} = \frac{1}{2Y_1}$$

$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$

back to PDF $f_{\underline{X}}(\underline{x}) = \begin{cases} 4X_1 X_2 & 0 < X_1 < 1, \ 0 < X_2 < 1 \\ 0 & \text{else} \end{cases}$

substitute $X_1 = \sqrt{Y_1 Y_2}$ $\quad X_2 = \sqrt{\frac{Y_2}{Y_1}}$ and use $|J|$

$$f_{\underline{Y}}(\underline{y}) = f_{\underline{X}}[h_1^{-1}(\underline{y}), h_2^{-1}(\underline{y})]\,|J|$$
$$= 4(\sqrt{Y_1 Y_2})\left(\sqrt{\frac{Y_2}{Y_1}}\right)\frac{1}{2Y_1}$$

$$f_{\underline{Y}}(\underline{y}) = \begin{cases} 2\frac{Y_2}{Y_1} & \text{for } (Y_1, Y_2) \in T \\ 0 & \text{else} \end{cases}$$

Lecture 8 (cont.) AMS 131

Go slow + keep track!

- start with PDF of $X$ vector $F_X(\underset{\sim}{X})$
- define the $Y$ vector: $Y_1 = h_1(X_1 \dots X_n)$ ... $Y_n = h_n(X_1 \dots X_n)$ — h functions, invertible
- inverse transform: $x_1 = h_1^{-1}(y_1 \dots y_n)$ ... $x_n = h_n^{-1}(y_1 \dots y_n)$
- evaluate PDF of $X$ at inverse image values

$$f_{\underset{\sim}{Y}}(\underset{\sim}{y}) = \begin{cases} f_{\underset{\sim}{X}}[h_1^{-1}(\underset{\sim}{y}), \dots, h_n^{-1}(\underset{\sim}{y})]\,|J| & \text{for } (y_1 \dots y_n) \in T \\ 0 & \text{else} \end{cases}$$

- work out Jacobian, matrix, absolute value
- work out image set T

$(X_1, X_2)$ joint dist. - you're only interested in the dist. of $Y_1 = h_1(X_1, X_2)$ ← univariate

1. find another R.V. $Y_2 = h_2(X_1, X_2)$ such that the transformation $(X_1, X_2) \to (Y_1, Y_2)$ is 1-1 with a differentiable inverse function, straight forward calcs
2. Work out joint dist. of $(Y_1, Y_2)$
3. integrate $Y_2$ out of joint dist, marginalize over $Y_2$, to get the marginal dist. of $Y_1$.

ex: $Y_2$ that wouldn't work 2D → 1D

$Y_1 = 2X_1$ $\quad$ $Y_2 = 3X_1 = \frac{3}{2}Y_1$ $\quad$ not invertible $Y_2 \sim Y_1$

$Y_2$ is linearly dependent on $Y_1$, so the rank of the 2x2 Jacobian matrix = 1, det(J) = 0

ex: earlier problem continued... $(X_1, X_2)$ joint continuous PDF

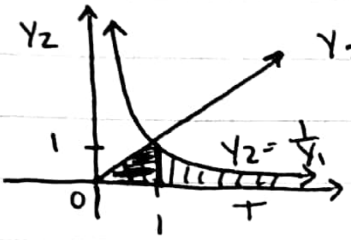
$$f_{X_1 X_2}(x_1, x_2) = \begin{cases} 4x_1x_2 & 0 < x_1, x_2 < 1 \\ 0 & \text{else} \end{cases}$$

$$(Y_1, Y_2) = \left(\frac{X_1}{X_2}, X_1X_2\right) \qquad f_{Y_1 Y_2}(y_1, y_2) = \begin{cases} 2\frac{y_2}{y_1} & (y_1, y_2) \in T \\ 0 & \text{else} \end{cases}$$ (PDF)

where $T = \{(y_1, y_2) : y_1 > 0,\ y_2 < \min(y_1, \frac{1}{y_1})\}$

if you only want marginal $Y_1$ → integrate $Y_2$ out of the joint distribution

AMS 131 Lecture 8 (cont.)

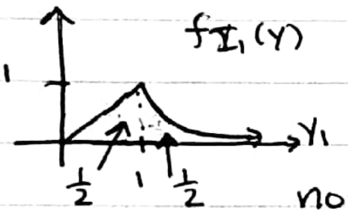



For $Y_1 > 0$, allowable $Y_2$ is:
- $0 < Y_1 < 1$, $0 < Y_2 < Y_1$
- $Y_1 > 1$, $0 < Y_2 < \frac{1}{Y_1}$

(below line vs. below hyperbola)

$$f_{Y_1}(y) \begin{cases} \int_0^{Y_1} 2\left(\frac{Y_2}{Y_1}\right) dY_2 = Y_1 & \text{for } 0 < Y_1 < 1 \\ \int_0^{\frac{1}{Y_1}} 2\left(\frac{Y_2}{Y_1}\right) dY_2 = Y_1^{-3} & \text{for } Y_1 > 1 \end{cases}$$

integrate $Y_2$ out




straight line into monotonically decreasing function.
not differentiable at $Y_1 = 1$, though

$\underset{\sim}{X}_n = (X_1, \dots, X_n)$ continuous with joint PDF $f_{X_1 \dots X_n}(X_1, \dots, X_n)$

$\underset{\sim}{Y}_n = (Y_1, \dots, Y_n)$ is a linear transformation of $\underset{\sim}{X}$:

$$\underset{\sim}{Y}^T = A \cdot \underset{\sim}{X}^T \leftarrow \text{transpose}$$

$(n \times 1) = (n \times n)(n \times 1)$

Where $A$ is an invertible, full rank matrix

$$\text{PDF } f_{\underset{\sim}{Y}}(\underset{\sim}{y}^T)_{n \times 1} = \frac{f_{\underset{\sim}{X}}(A^{-1}\underset{\sim}{y}^T)}{|\det(A)|} \leftarrow (n \times n)(n \times 1)$$

$Y_1 = X_1 + X_2$

$Y_2 = X_1 - X_2$

$$\underset{2 \times 2}{A} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \det(A) = -2 \quad |\det(A)| = 2$$

$$A^{-1} = \frac{1}{-2}\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} = \frac{1}{2}A \qquad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$A^{-1} = \frac{\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}}{(ad - bc)}$$

$\det(A) = ad - bc$