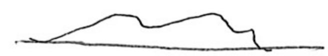
STAT 24400
Lecture 4
January 14 2016

Describing Probability Distributions

of random vars

Full description: $p_X(x)$ or $f(x)$



Partial description (much shorter!):

Measures of center and dispersion

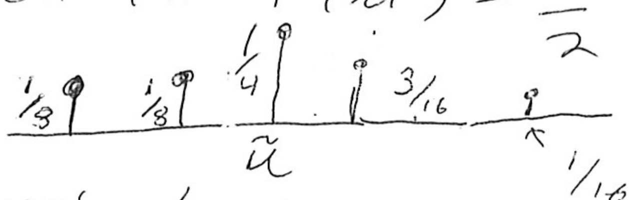
Center: Expectation ("expected value" "mean")

Median (implies "middle")

First, consider Median:

$P(X \geq \tilde{\mu}) \geq \frac{1}{2}$ $\Leftarrow$ $\tilde{\mu}$ such that $F(\tilde{\mu}) = \frac{1}{2}$

$P(X \leq \tilde{\mu}) \leq 1/2$



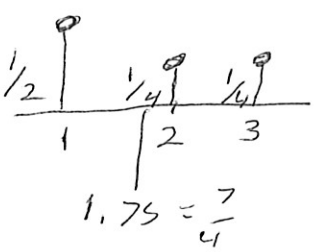
May not be unique!

Expectation of X

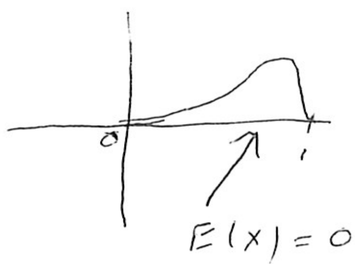
"Center of Gravity", 'Fair Price'

Def $E(X) = \begin{cases} \sum_{\text{all } x} x \, p_X(x) & \text{Discrete} \\ \int_{-\infty}^{\infty} x \, f_X(x)\, dx & \text{continuous} \end{cases}$

Simple Examples:



$$E(X) = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{4} = \frac{7}{4}$$



$$f_X(x) = \begin{cases} 12x^2(1-x) & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$E(X) = \int_{-\infty}^{\infty} x \, f_X(x) dx = \int_0^1 x \; 12x^2(1-x)\, dx$$

$$= 12 \int_0^1 (x^3 - x^4)\, dx = 12 \left[\frac{x^4}{4} - \frac{x^5}{5} \right|_0^1$$

$$= \frac{12}{20} = 0.6$$

$$E(X) = \begin{cases} \sum_{\text{all } x} x\, p_X(x) & \underline{\text{discrete}} \\ \int_{-\infty}^{\infty} x\, f_X(x)\,dx & \underline{\text{continuous}} \end{cases}$$

So... suppose we have a function of a random variable $h(X)$.

What is $E(X^2)$? $E(e^X)$? $E(\ln X)$?

Is $E(h(X)) = h(E(X))$? (NO, usually NOT)

$Y = h(X)$. 2 ways to find $E(Y)$.

① Find $f_Y(y)$, $E(Y) = \int_{-\infty}^{\infty} y\, f_Y(y)\,dy$

② $E(Y) = \int_{-\infty}^{\infty} h(x)\, f_X(x)\,dx$ (discrete: $E(Y) = \sum h(x)\, p_X(x)$)

Usually ② is easier.

Pf $E(Y) = \int_{-\infty}^{\infty} y\, f_Y(y)\,dy = \int_{-\infty}^{\infty} y\, f_X(g(y))\,g'(y)\,dy$

but $x = g(y)$ $y = h(x)$ $dx = g'(y)\,dy$

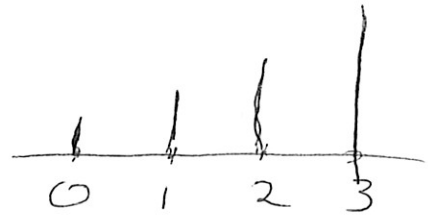
so:

$$E(Y) = \int_{-\infty}^{\infty} h(x)\, f_X(x)\,dx$$

Examples

①

X	0	1	2	3
$p_X(x)$	.1	.2	.3	.4

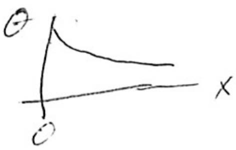


$$E(X) = 0(0.1) + 1(0.2) + 2(0.3) + 3(0.4)$$
$$= \boxed{2}$$

$$E(X^2) = 0^2(0.1) + 1^2(0.2) + 2^2(0.3) + 3^2(0.4)$$
$$= \boxed{5}$$

NOTE: $E(X^2) \neq (E(X))^2$

② X exponential, $f_X(x) = \theta e^{-\theta x}, \; x > 0$



$$E(X) = \int_0^\infty x\,\theta e^{-\theta x}\,dx = \frac{1}{\theta}\int_0^\infty u e^{-u}\,du$$ (ch. vars: $u = \theta x$, $\frac{1}{\theta}du = dx$)

$$\int_0^\infty u e^{-u} = -ue^{-u}\Big|_0^\infty + \int_0^\infty e^{-u}\,du$$ (integration by parts)
$$= 0 + -e^{-u}\Big|_0^\infty = 1$$

So $E(X) = \left(\frac{1}{\theta}\right)(1) = \frac{1}{\theta}$

$$E(X^2) = \int_0^\infty x^2\theta e^{-\theta x}\,dx = \frac{1}{\theta^2}\int_0^\infty u^2 e^{-u}\,du$$

$$\int_0^\infty u^2 e^{-u}\,du = -u^2e^{-u}\Big|_0^\infty + \int_0^\infty 2ue^{-u}\,du$$
$$= 2\int_0^\infty ue^{-u}\,du = 2$$ (see above)

So $E(X^2) = \frac{2}{\theta^2} \neq E(X)$

General Examples

① Linear Transformations $\quad h(x) = aX + b$

Then $\quad E(aX+b) = aE(X) + b$

(a case where $E(h(X)) = h(E(X))$)

Pf $\quad E(aX+b) = \int_{-\infty}^{\infty} (ax+b)\, f_X(x)\,dx$

$$= \int_{-\infty}^{\infty} \left(ax f_X(x) + b f_X(x)\right) dx = a\underbrace{\int_{-\infty}^{\infty} x f_X(x)\,dx}_{E(X)} + b\underbrace{\int_{-\infty}^{\infty} f_X(x)\,dx}_{1}$$

$$\boxed{= aE(X) + b} \quad \therefore$$

Note that if $a = 0$,
get $E(b) = b$,
True for all constants including $E(X)$

$E(E(X)) = E(X)$

② Variances $\quad h(X) = (X - \mu_X)^2$

Def: $Var(X) = E[(X-\mu_X)^2]$ is variance of X (or of prob dist of X)

Notation: $Var(X) = \sigma_X^2$ (or just σ^2 when no confusion)

$\sqrt{Var(X)} = \sigma_X$ the standard deviation of X

var and s.d. measure dispersion, spread.

For Theoretical Calculation:

$$\boxed{Var(X) = E(X^2) - \mu_X^2}$$

pf: $Var(X) = E[(X-\mu_X)^2]$

$$= \int_{-\infty}^{\infty} (x^2 - 2x\mu_X + \mu_X^2) f_X(x)\,dx$$

$$= \int_{-\infty}^{\infty} x^2 f_X(x)\,dx - 2\mu_X \int_{-\infty}^{\infty} x f_X(x)\,dx + \mu_X^2 \int_{-\infty}^{\infty} f_X(x)\,dx$$

$$= E(X^2) - 2\mu_X \cdot \mu_X + \mu_X^2 \cdot 1$$

$$= E(X^2) - \mu_X^2$$

Examples: Exponential dist

$$E(X^2) = \frac{2}{\theta^2}, \quad E(X) = \frac{1}{\theta}$$

$$Var(X) = \frac{2}{\theta^2} - \frac{1}{\theta^2} = \boxed{\frac{1}{\theta^2}}$$

Linear Transformation:

$$E(aX+b) = a\mu_X + b$$

$$Var(aX+b) = a^2 Var(X)$$

(so $\mu_{aX+b} = a\mu_X + b$; $\sigma_{aX+b}^2 = a^2\sigma_X^2$)

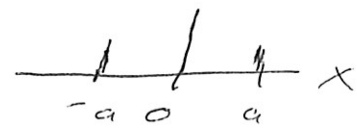
Special Case: $W = \frac{X - \mu_X}{\sigma_X}$ // Standard Form

$$E(W) = \frac{E(X) - \mu_X}{\sigma_X} = 0; \quad Var(W) = \sigma_W^2 = 1$$

Interpreting Variance

example:

x	$-a$	0	a
$p_X(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$



$E(X) = 0$ (by symmetry, or $-a(\frac{1}{4}) + 0(\frac{1}{2}) + a(\frac{1}{4}) = 0$)

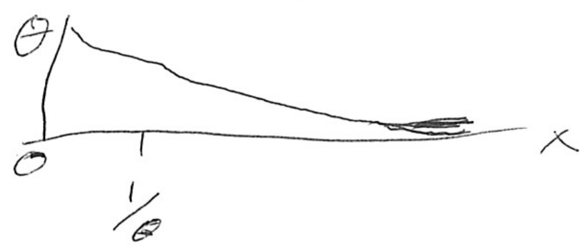
so

$$Var(X) = E(X^2)$$
$$= (-a)^2(\tfrac{1}{4}) + 0^2(\tfrac{1}{2}) + a^2(\tfrac{1}{4})$$
$$= \frac{a^2}{2}$$

large $a \Leftrightarrow$ large $Var(X) \Leftrightarrow$ large spread

note: squared units!

Exponential:



$$E(X) = \frac{1}{\theta}$$

$$Var(X) = \frac{1}{\theta^2}, \quad \sigma_X = \frac{1}{\theta}$$

Multivariate, or Joint Distributions

Distributions of 2, 3, 4, or more random vars.

Discrete Bivariate Case:

X, Y are 2 rand. vars. defined on same sample space.

Bivariate prob. function:

$$p(x,y) = Pr(X = x \text{ and } Y = y)$$

$$\sum_{\text{all } x} \sum_{\text{all } y} p(x,y) = 1$$

Ex: (1) Toss 2 fair coins 3 times each

$X = \#H\text{'s}$ coin 1

$Y = \#T\text{'s}$ coin 2

$Z = \#T\text{'s}$ coin 1

dist of (X, Y)

$x \backslash y$	0	1	2	3
0	$1/64$	$3/64$	$3/64$	$1/64$
1	$3/64$	$9/64$	$9/64$	$3/64$
2	$3/64$	$9/64$	$9/64$	$3/64$
3	$1/64$	$3/64$	$3/64$	$3/64$

dist of (X, Z)

$x \backslash z$	0	1	2	3
0	0	0	0	$1/8$
1	0	0	$3/8$	0
2	0	$3/8$	0	0
3	$1/8$	0	0	0

Can compute univariate prob. dist. from bivariate distributions by addition:

$$p_X(x) = \sum_{\text{all } y} p(x,y)$$
$$p_Y(y) = \sum_{\text{all } x} p(x,y)$$

These are called the marginal prob functions of X and Y respectively.

Idea: The event

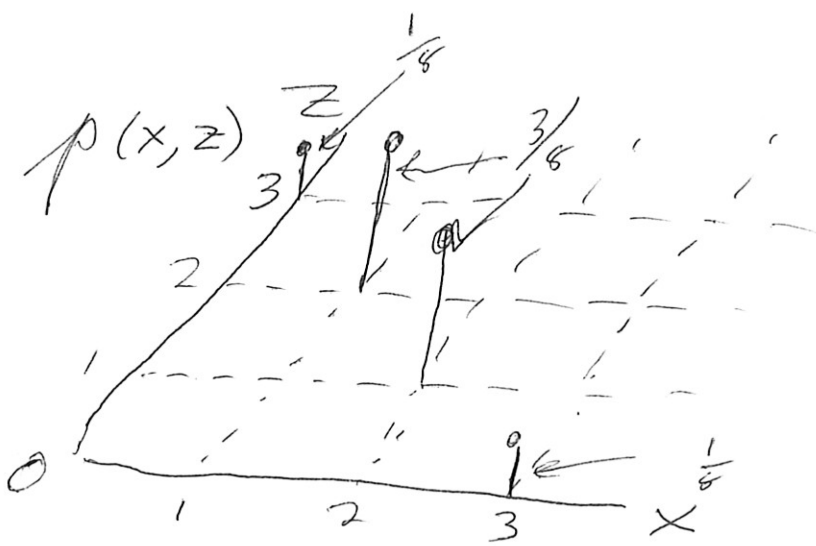
$$\{X=x\} = \{X=x \text{ and } Y=1\} \cup \{X=x \text{ and } Y=2\} \dots$$
$$\dots \cup \{X=x \text{ and } Y=57\} \dots$$

The events on the right side are mutually exclusive, so we can add probabilities

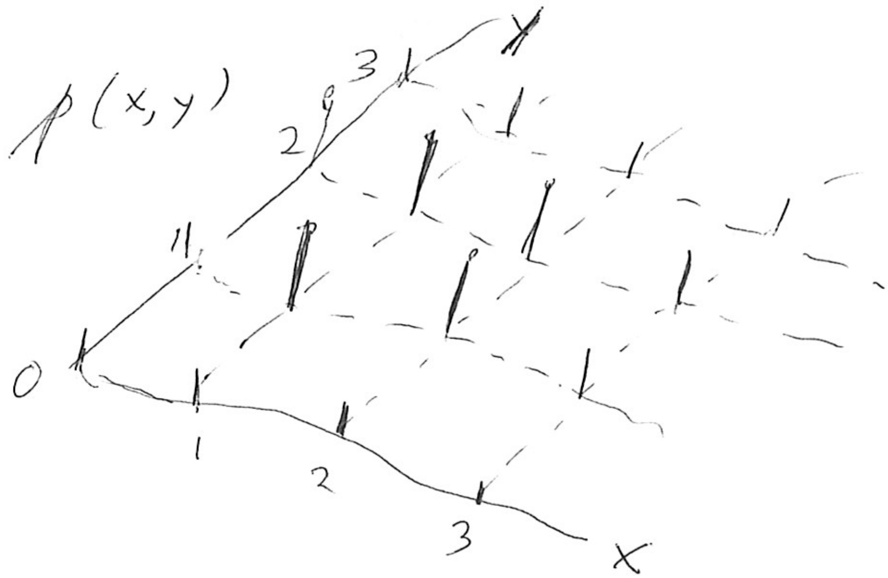
3 coin example, part 2:

X \ Y	0	1	2	3	
0	1/64	3/64	3/64	1/64	1/8
1	3/64	9/64	9/64	3/64	3/8
2	3/64	9/64	9/64	3/64	3/8
3	1/64	3/64	3/64	1/64	1/8
	1/8	3/8	3/8	1/8	1

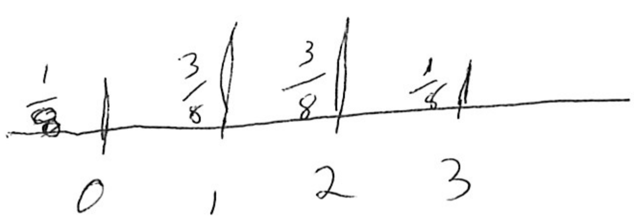
X \ Z	0	1	2	3	
0	0	0	0	1/8	1/8
1	0	0	3/8	0	3/8
2	0	3/8	0	0	3/8
3	1/8	0	0	0	1/8
	1/8	3/8	3/8	1/8	1



Z = 3 − X
not indep!



X, Y
indep.



marginal dist
is like a
"side view"
of joint dist.

IMPORTANT

Conditional Prob. Functions

$$p(y/x) = Pr(\text{"}Y=y\text{"} \mid \text{"}X=x\text{"})$$ "given"

$$= \frac{Pr(\text{"}X=x\text{" AND "}Y=y\text{"})}{Pr(\text{"}X=x\text{"})}$$

$$p(y/x) = \frac{p(x,y)}{p_x(x)}, \quad \sum_{\text{all } y} p(y/x) = 1$$

If $p(y/x) = p_Y(y)$ for all x, y

(ie $p(y/x) = \frac{p(x,y)}{p_X(x)}$, so $p(x,y) = p_X(x)\,p_Y(y)$)

we say the random variables X and Y are <u>independent</u>

In the previous example

X and Y are <u>indep</u>

X and Z are <u>dependent</u>

This is true even though they have the <u>same</u> marginal distribution.

Discrete Example:

$X = \#H$'s
$Z = \#T$'s] Same coin

$X + Z = 3$

$X \backslash Z$	0	1	2	3
0	0	0	0	$\frac{1}{8}$
1	0	0	$\frac{3}{8}$	0
2	0	$\frac{3}{8}$	0	0
3	$\frac{1}{8}$	0	0	0
	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

"Marginal" $p_Z(z)$:

z	0	1	2	3
p	$1/8$	$3/8$	$3/8$	$1/8$

"Conditional" $p(x|z)$, say for $Z=2$:

$$p(x|Z=2) = \frac{p(x,2)}{p_Z(2)}$$

x	0	1	2	3
$p(x\|Z=2)$	0	1	0	0

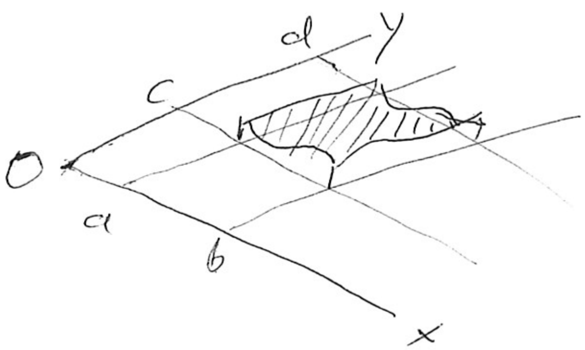
Continuous bivariate case

Bivariate prob. density:

① $f(x, y) \geq 0$

② $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\,dx\,dy = 1$ (a double integral)

$$\Pr(a < X < b \text{ and } c < Y < d) = \underbrace{\int_c^d \int_a^b f(x, y)\,dx\,dy}_{\text{volume}}$$



volume between the x-y plane and the surface $f(x, y)$

Marginal prob. densities

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\,dy\,; \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\,dx$$

Conditional prob densities

$$f(y|x) = \frac{f(x, y)}{f_X(x)}$$

IMPORTANT

If $f(y|x) = f_Y(y)$ for all x, y,
X and Y are independent and $f(x, y) = f_X(x)\, f_Y(y)$

Interpreting Conditional Densities

Intuitively, think of

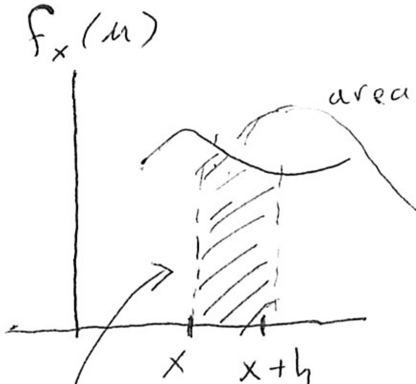
$$f(y|x)\,dy = \Pr(y \le Y \le y+dy \mid X = x)$$

even though the right hand side is not defined (since $\Pr(X=x)=0$).

Under regularity conditions, one can get (for h small)

$$\Pr(a < Y \le b \mid X=x) \cong \Pr(a < Y \le b \mid x \le X \le x+h)$$

$$= \frac{\Pr(a<Y\le b \text{ AND } x \le X \le x+h)}{\Pr(x \le X \le x+h)}$$



$$= \frac{\int_a^b \int_x^{x+h} f(u,y)\,du\,dy}{\int_x^{x+h} f_X(u)\,du}$$

almost a rectangle with base h

$$\cong \frac{\int_a^b f(x,y)\,h\,dy}{f_X(x)\,h}$$

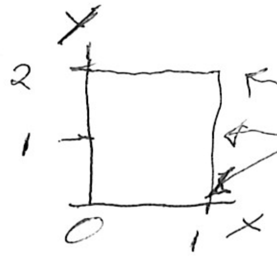
$$= \int_a^b \left[\frac{f(x,y)}{f_X(x)}\right] dy$$

conditional probability by area

Example X, Y random variables

joint (bivariate) density

$$f(x,y) = \begin{cases} y\left(\frac{1}{2} - x\right) + x & \text{for } 0<x<1 \\ & \phantom{\text{for }} 0<y<2 \\ 0 & \text{otherwise} \end{cases}$$

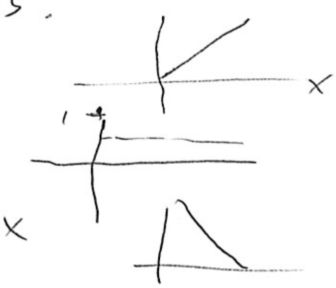


cross-sections:

$f(x,0) = x$

$f(x,1) = \frac{1}{2}$

$f(x,2) = 1 - x$



$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = \iint \left[y\left(\tfrac{1}{2}-x\right)+x\right] dx\,dy$$

$$= \int_0^2 \left[\int_0^1 \left[y\left(\tfrac{1}{2}-x\right)+x\right] dx\right] dy$$

$$= \int_0^2 \left[\int_0^1 y\left(\tfrac{1}{2}-x\right) dx + \int_0^1 x\,dx\right] dy$$

$$= \int_0^2 \left[y \int_0^1 \left(\tfrac{1}{2}-x\right) dx + \int_0^1 x\,dx\right] dy$$

$$= \int_0^2 \left[y \cdot 0 + \frac{1}{2}\right] dy$$

$$= \int_0^2 \frac{1}{2}\,dy = \frac{1}{2} \cdot 2 = 1$$

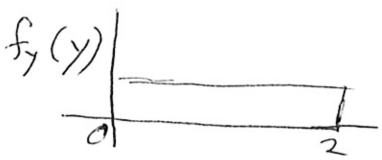
In the process, we found

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^1 \left[y\left(\tfrac{1}{2}-x\right)+x\right]dx = \tfrac{1}{2}$$

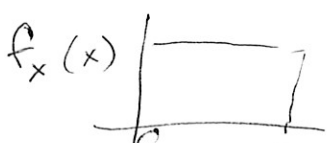
$\rightarrow$ This is $f_Y(y) = \begin{cases} \frac{1}{2} & 0 \le y \le 2 \\ 0 & \text{otherwise} \end{cases}$

ie, the marginal of Y is <u>uniform</u>.

Similarly $f_X(x) = \int_0^2 \left(y\left(\tfrac{1}{2}-x\right)+x\right)dy$



$$= \frac{y^2}{2}\Big|_0^2 \left(\tfrac{1}{2}-x\right) + 2x = 1 + 2x - 2x$$

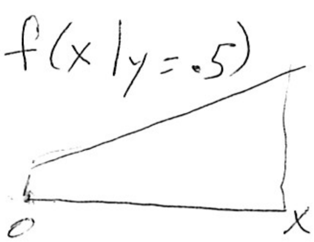


$$= \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

also uniform!

but X, Y not indep:

$$f(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{y\left(\frac{1}{2}-x\right)+x}{\left(\frac{1}{2}\right)}$$



$$= \begin{cases} y(1-2x)+2x & 0 \le x \le 1 \\ 0 & \end{cases}$$

depends on y!

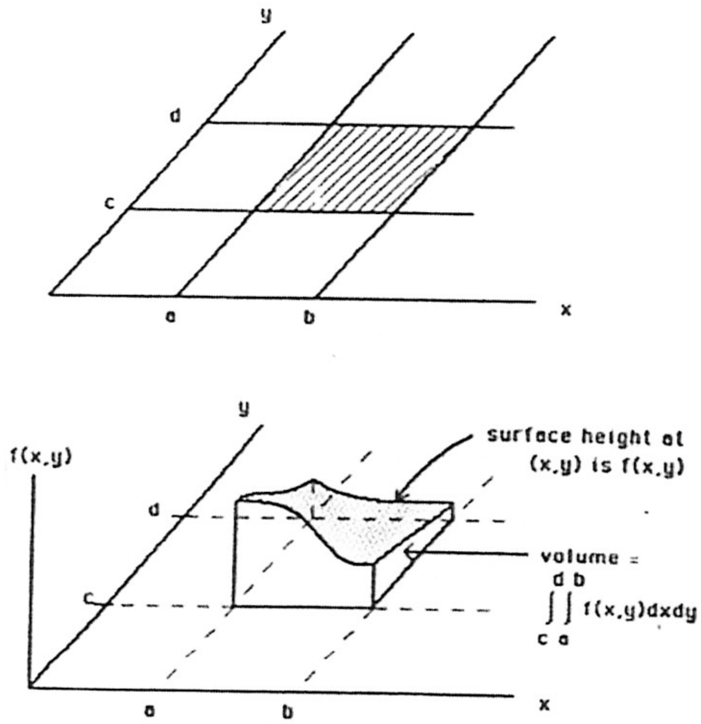


Figure 3.3. A bivariate density.

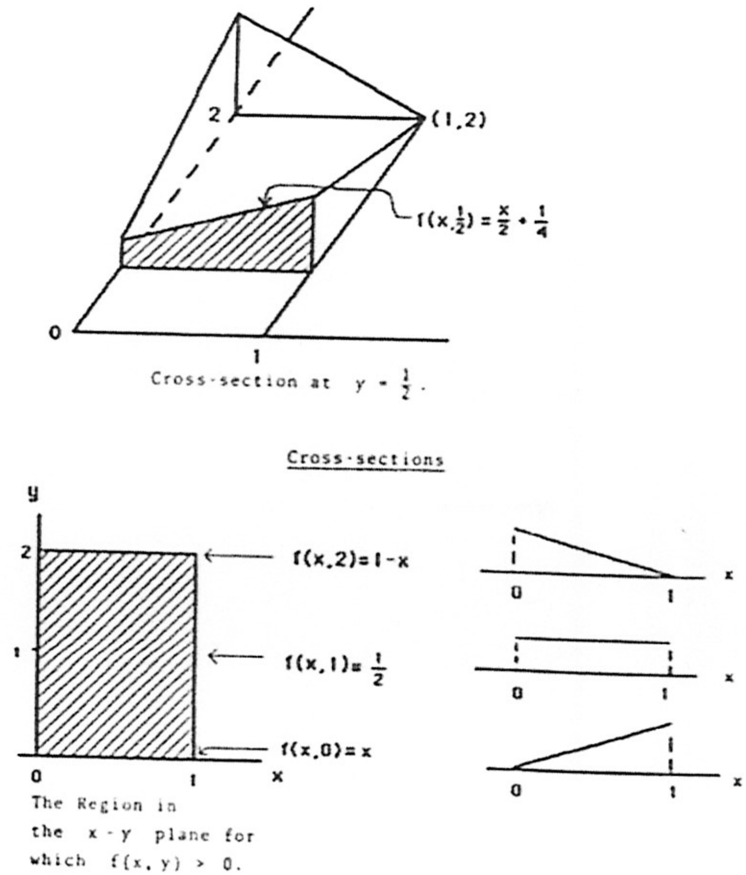


Figure 3.4. A bivariate density and its cross-sections.