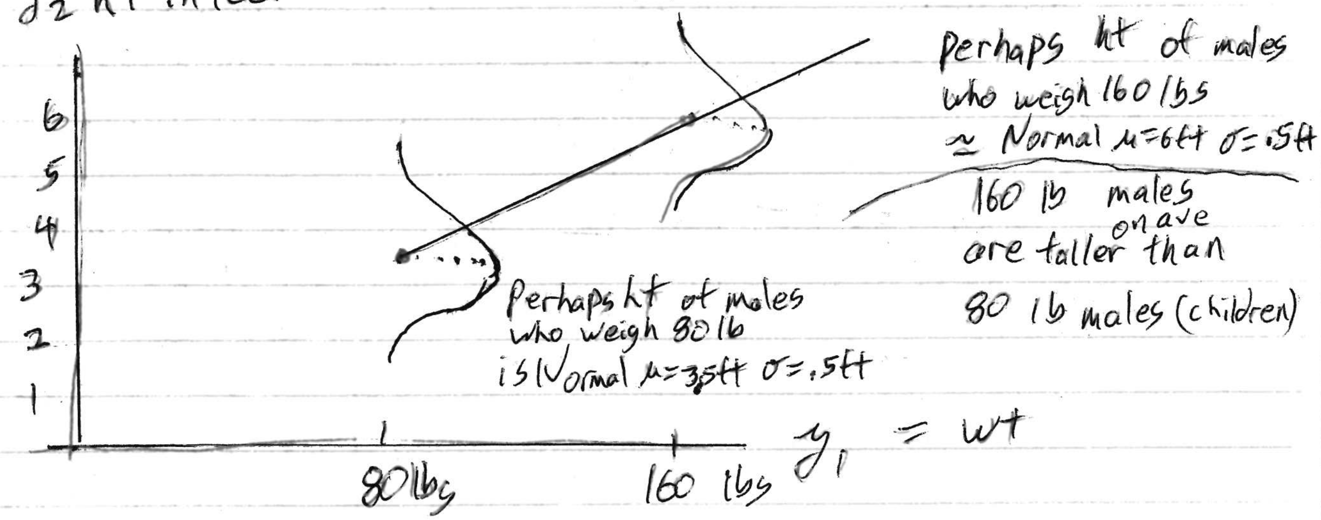
M483 30

19] p240 $F(y_1|y_2) = P(Y_1 \le y_1 | Y_2 = y_2)$

$$= \int_{-\infty}^{y_1} \underset{Y_1|Y_2=y_2}{f(t_1|y_2)}\, dt_1$$



ex] $f(y_1,y_2) = \begin{cases} \frac{6}{5}(y_1 + y_2^2), & 0 \le y_1 \le 1, 0 \le y_2 \le 1 \\ 0, & \text{else} \end{cases}$

Find $f_{Y_1}(y_1)$, $\underset{Y_2|Y_1=.8}{f(y_2|.8)}$ and $P(Y_2 \le .5 | Y_1 = .8)$

Soln) $f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f(y_1,y_2)\,dy_2 = \int_0^1 \frac{6}{5}(y_1+y_2^2)\,dy_2$

$$= \frac{6}{5}\left(y_1 y_2 + \frac{y_2^3}{3}\right)\Big|_0^1 = \frac{6}{5}\left(\frac{1}{3} + y_1\right), 0<y_1<1$$

so $f_{Y_2|Y_1=.8}(y_2|.8) = \frac{f(.8,y_2)}{f_{Y_1}(.8)} = \frac{\frac{6}{5}(0.8+y_2^2)}{\frac{6}{5}(\frac{1}{3}+.8)} = \frac{0.8+y_2^2}{\frac{1}{3}+\frac{8}{10}}$

$$= \frac{30}{34}(.8+y_2^2) = \frac{24+30y_2^2}{34}, \quad 0<y_2<1$$

$P(Y_2 \le .5 | Y_1 = .8) = \int_0^{.5} \frac{24+30y_2^2}{34}\,dy_2 = \frac{24}{34}y_2 + \frac{30}{34}\frac{y_2^3}{3}\Big|_0^{.5}$

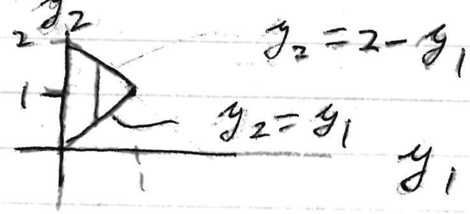
$= \frac{12}{34} + \frac{10}{34}\left(\frac{1}{8}\right) = \frac{106}{34(8)} = .3897$

30.5

20) With conditional probabilities, the domain can depend on the other variable if the region of integration is not rectangular

ex) problem 5.32 on p 245

$$f(y_1, y_2) = \begin{cases} 6\, y_1^2\, y_2, & 0 \le y_1 \le y_2 \\ & y_1 + y_2 \le 2 \\ 0 & \text{else} \end{cases}$$



a) $f_{Y_1}(y_1) = \int_{y_1}^{2-y_1} 6 y_1^2\, y_2\, dy_2 = 6 y_1^2 \frac{y_2^2}{2} \Big|_{y_1}^{2-y_1}$

$= 6 y_1^2 \dfrac{(2-y_1)^2 - y_1^2}{2}$

$= 3 y_1^2 \left(4 - 4y_1 + y_1^2 - y_1^2\right)$

$= 12\, y_1^2 (1 - y_1), \quad 0 \le y_1 \le 1$

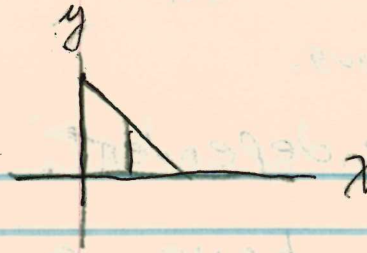
c) $f_{Y_2 | Y_1 = y_1}(y_2 | y_1) = \dfrac{f(y_1, y_2)}{f_{Y_1}(y_1)} = \dfrac{6\, y_1^2\, y_2}{12 y_1^2 (1 - y_1)}$

$= \dfrac{y_2}{2(1-y_1)}$ for $\underbrace{y_1 < y_2 < 2 - y_1}_{\text{important}}$

if $0 < y_1 < 1$.

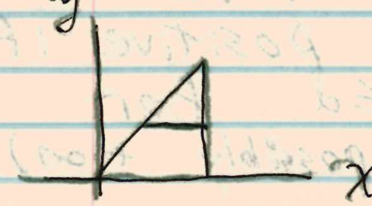
See HW 10 E c, d.

Regions of integration



$$\int_{\text{function of } x}^{\text{function of } x} h(x, y)\, dy$$

dy means line is parallel to the y axis



$$\int_{\text{function of } y}^{\text{function of } y} h(x, y)\, dx$$

dx means line is parallel to the x axis

y_2 vertical

y_1 horizontal

The text uses y_2 for y and y_1 for x.

endExam 2 material

§5.4 21] know Random variables Y_1 and Y_2 are independent if any one of the following conditions holds

i) $F(y_1, y_2) = F_{Y_1}(y_1)\, F_{Y_2}(y_2) \quad \forall\, y_1, y_2$

ii) $p(y_1, y_2) = p_{Y_1}(y_1)\, p_{Y_2}(y_2) \quad \forall\, y_1, y_2$ where Y_1, Y_2 are discrete

iii) $f(y_1, y_2) = f_{Y_1}(y_1)\, f_{Y_2}(y_2) \quad \forall\, y_1, y_2$ where Y_1, Y_2 are continuous

Otherwise Y_1 and Y_2 are dependent.

22) know p 251 $Y_1, Y_2, \ldots, Y_n$ are independent if $F(y_1, y_2, \ldots, y_n) = F_{Y_1}(y_1)\, F_{Y_2}(y_2) \cdots F_{Y_n}(y_n)$

or if $p(y_1, y_2, \ldots, y_n) = p_{Y_1}(y_1)\, p_{Y_2}(y_2) \cdots p_{Y_n}(y_n)$ where $Y_1, \ldots, Y_n$ are discrete or if $f(y_1, \ldots, y_n) = f_{Y_1}(y_1) \cdots f_{Y_n}(y_n)$

if $Y_1, \ldots, Y_n$ are continuous.

Otherwise $Y_1, \ldots, Y_n$ are dependent.

23] * P250 Let Y_1 and Y_2 have a joint density $f(y_1, y_2)$ that is positive iff

support of Y_1 and Y_2 $\Big\{$ $a \le y_1 \le b$ and $c \le y_2 \le d$ for constants a, b, c, d (possibly $\pm\infty$).

Then Y_1 and Y_2 are independent iff $f(y_1, y_2) = g(y_1)\, h(y_2)$ where g

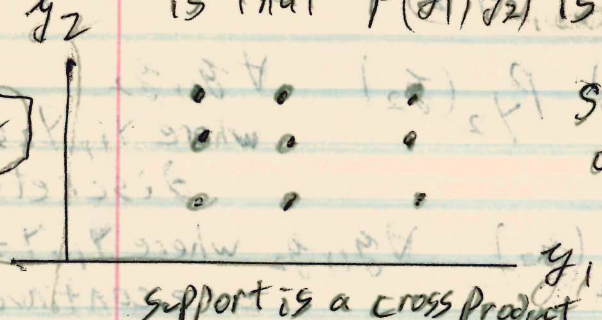
is a nonnegative function of y_1 alone and h is a nonnegative function of y_2 alone.

Note] The support $\mathcal{Y}_i$ of $Y_i = \{y : f(y) \text{ or } P(y) > 0\}$. The support $\mathcal{Y}$ of Y_1 and $Y_2 = \{(y_1, y_2) : f(y_1, y_2) > 0 \text{ or } P(y_1, y_2) > 0\}$. The support $\mathcal{Y}$ is a cross product if $\mathcal{Y} = \mathcal{Y}_1 \times \mathcal{Y}_2 = \{(y_1, y_2) : y_1 \in \mathcal{Y}_1 \text{ and } y_2 \in \mathcal{Y}_2\}$.

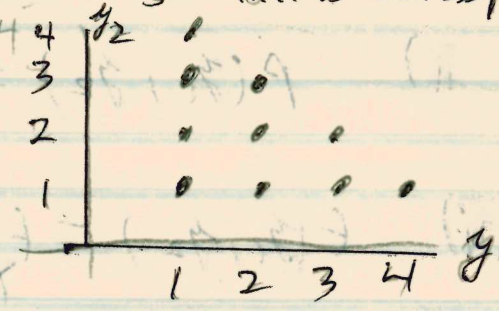
ex] The support is rectangular if $\mathcal{Y}_1$ and $\mathcal{Y}_2$ are intervals eg $\mathcal{Y} = \{(y_1, y_2) : y_1 \in (a, \infty) \text{ and } y_2 \in (c, d)\}$.

A necessary condition (but not sufficient) for 2 discrete RVs to be independent is that $P(y_1, y_2)$ is positive on a grid that is a cross product.

ex]



Support of Y_1 and Y_2



support is a cross product

Y_1 and Y_2 could be independent or dependent

Y_1 and Y_2 are dependent

for example if $Y_1 = 4$, then $Y_2 = 1$

Note The reasoning for continuous RV's is similar.

If Y_1 is near 1 then Y_2 is near 0 so dependent.

24] If Y_1 and Y_2 are discrete RVs, 483 32

the set $\{(y_1, y_2) : p(y_1, y_2) > 0\}$

p313 is called the support of (Y_1, Y_2).

cross product support

If Y_1 and Y_2 are continuous,

$\{(y_1, y_2) : f(y_1, y_2) > 0\}$ is the support of (Y_1, Y_2).

For independence the support needs to be a cross product.

ex] $f(y_1, y_2) = \begin{cases} 6\ (1-y_2) & 0 < y_1 < y_2 < 1 \\ 0 & \text{else} \end{cases}$

$Y_1 = y_2$

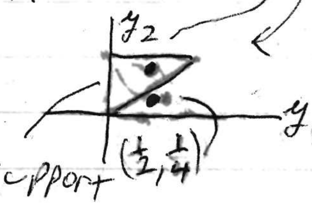
$Y_2 = Y_1$

Let $g(y_1) = 6$ for $0 < y_1 < 1$

and $h(y_2) = 1 - y_2$ for $0 < y_2 \leq 1$.

Are Y_1 and Y_2 independent? $\left(\frac{1}{2}, \frac{3}{4}\right)$

No, support is not a cross product.



Also $f\left(\frac{1}{2}, \frac{1}{4}\right) = 0 \neq f_1\left(\frac{1}{2}\right) f_2\left(\frac{1}{4}\right)$.

25] Independence is the assumption that makes multiple integration tractable since the region of integration is a cross product $\mathcal{Y} = \mathcal{Y}_1 \times \cdots \times \mathcal{Y}_n$. Since if $Y_1, \ldots, Y_n$ are independent, then

$$\int_{a_n}^{b_n} \int_{a_{n-1}}^{b_{n-1}} \cdots \int_{a_1}^{b_1} f(y_1, \ldots, y_n)\, dy_1 \cdots dy_n$$

$$= \int_{a_1}^{b_1} f_{Y_1}(y_1)\, dy_1 \int_{a_2}^{b_2} f_{Y_2}(y_2)\, dy_2 \cdots \int_{a_n}^{b_n} f_{Y_n}(y_n)\, dy_n$$

and $\int_{a_n}^{b_n} \cdots \int_{a_1}^{b_1} \left(\prod_{i=1}^{n} g(y_i)\right) f(y_1, \ldots, y_n)\, dy_1 \cdots dy_n = \prod_{i=1}^{n} \int_{a_i}^{b_i} g(y_i) f_{Y_i}(y_i)\, dy_i$

where $\prod_{i=1}^{n} d_i = \underbrace{d_1 \cdot d_2 \cdots d_n}_{\text{multiply}}$ 32.5

26] To show dependence, find a single point $(y_1, \dots, y_n)$ such that

$$f(y_1, \dots, y_n) \neq f_{Y_1}(y_1) \cdots f_{Y_n}(y_n).$$

In problems, usually $n=2$.

27] If $p(y_1, y_2)$ is given by a table, Y_1 rows Y_2 columns, independence holds if (ith row sum)(jth col sum) = ij table entry for all ij entries.

ex)

$Y_1 \backslash Y_2$	1	2	3	4	
1	.06	.02	.04	.08	.20
2	.15	.05	.10	.20	.50
3	.09	.03	.06	.12	.30
	.30	.10	.20	.40	

If Y_1 and Y_2 are independent, have to check 12 products:

.3(.2)	.1(.2)	.2(.2)	.4(.2)
.3(.5)	.1(.5)	.2(.5)	.4(.5)
.3(.3)	.1(.3)	.2(.3)	.4(.3)

Get Y_1 and Y_2 are independent

Note Dependence is easier: as soon as one pair (y_1, y_2) is found such that $p(y_1, y_2) \neq p_{Y_1}(y_1)\,p_{Y_2}(y_2)$ we can claim that Y_1 and Y_2 are dependent.

See ex 5.10 on p 248.

§5.5 p 256 If $p(y_1,\dots,y_K)$ is the joint prob function, then

28]* $E\,g(Y_1,\dots,Y_K) = \sum_{y_K}\sum_{y_{K-1}}\cdots\sum_{y_1} g(y_1,\dots,y_K)\,p(y_1,\dots,y_K)$

29]* If $f(y_1,\dots,y_K)$ is the joint pdf, then

$$E\,g(Y_1,\dots,Y_K) = \int_{a_{y_K}}\int_{a_{y_{K-1}}}\cdots\int_{a_{y_2}}\int_{a_{y_1}} g(y_1,\dots,y_K)\,f(y_1,\dots,y_K)\,dy_1\cdots dy_K$$

where a_{y_i} are the limits of integration for dy_i.

Note: Usually $K=2$ and $g(y_1,y_2) = y_1^i\,y_2^j$

for nonnegative integers i and j.

$E[Y_1]$, $E[Y_2]$, $E[Y_1Y_2]$ are especially common.

Hard way:

30] $E(Y_1) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_1\,f(y_1,y_2)\,dy_2\,dy_1$

$= \int_{-\infty}^{\infty} y_1\left[\int_{-\infty}^{\infty} f(y_1,y_2)\,dy_2\right]dy_1$

$= \int_{-\infty}^{\infty} y_1\,f_{Y_1}(y_1)\,dy_1$ as before.

easier In general, $E\,g(Y_1) = \int_{-\infty}^{\infty} g(y_1)\,f_{Y_1}(y_1)\,dy_1$

as before. Similarly,

$E\,g(Y_2) = \int_{-\infty}^{\infty} g(y_2)\,f_{Y_2}(y_2)\,dy_2$

Hard way

31] $E\,g(Y_1) = \sum_{y_1}\sum_{y_2} g(y_1)\,p(y_1,y_2) =$

$\sum_{y_1}\left[g(y_1)\sum_{y_2} p(y_1,y_2)\right] = \sum g(y_1)\,p_{Y_1}(y_1)$ as before.

§5.6 p258 32] $E(c) = c$ for any constant c 33.5

33] * $E[g_1(Y_1, Y_2) + g_2(Y_1, Y_2) + \cdots + g_k(Y_1, Y_2)]$

$= E[g_1(Y_1, Y_2)] + E[g_2(Y_1, Y_2)] + \cdots + E[g_k(Y_1, Y_2)]$

$\boxed{\text{In particular}}$, $E(aY_1 + bY_2) = aE(Y_1) + bE(Y_2)$

34] * p259 Let Y_1 and Y_2 be independent and $g(Y_1)$ and $h(Y_2)$ functions of only Y_1 and Y_2, respectively. Then $\underbrace{E[g(Y_1)h(Y_2)]}_{\text{joint}} = \underbrace{E[g(Y_1)]\,E[h(Y_2)]}_{\text{univariate}}$ provided that the expectations exist.

$\boxed{\text{In particular}}$, $E(Y_1 Y_2) = E(Y_1)E(Y_2)$ if $Y_1 \perp\!\!\!\perp Y_2$ ind

proof for discrete case $E[g(Y_1)h(Y_2)] =$

$\sum_{y_1}\sum_{y_2} g(y_1)h(y_2)\,p(y_1, y_2) \underset{\text{ind}}{=}$

$\sum_{y_1}\sum_{y_2} g(y_1)h(y_2)\,p_{Y_1}(y_1)\,p_{Y_2}(y_2)$

$= \sum_{y_1} g(y_1)\,p_{Y_1}(y_1) \sum_{y_2} h(y_2)\,p_{Y_2}(y_2) = [Eg(Y_1)][Eh(Y_2)]$

§5.7 p265 35] Know If Y_1 and Y_2 are RV's, then the covariance of Y_1 and Y_2 is

$\text{cov}(Y_1, Y_2) = E[(Y_1 - E(Y_1))(Y_2 - E(Y_2))]$.

36] know p266 Short cut formula

$\text{Cov}(Y_1, Y_2) = E(Y_1 Y_2) - E(Y_1)E(Y_2)$

Note: $\text{cov}(Y_1, Y_1) = \text{Var}\, Y_1$

37] * p265 The <u>correlation coefficient</u> $\rho = \dfrac{\operatorname{cov}(Y_1, Y_2)}{\sqrt{V(Y_1)\;V(Y_2)}}$

p265 ρ is a measure of the linear dependence of Y_1 and Y_2

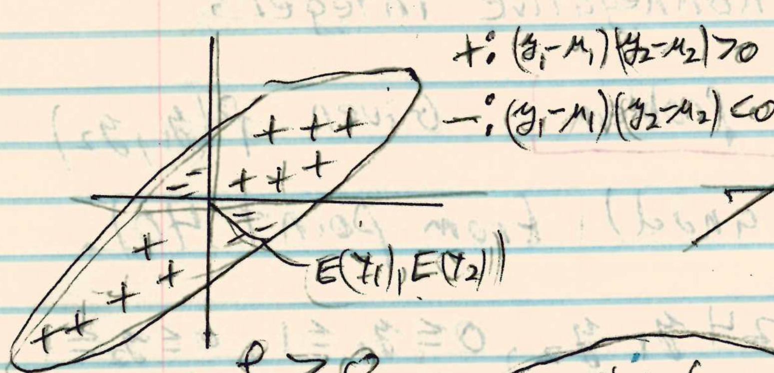
38] properties of ρ: $\rho 1)$ $-1 \leq \rho \leq 1$

$\rho 2)$ $\rho > 0$ means Y_2 increases as Y_1 increases

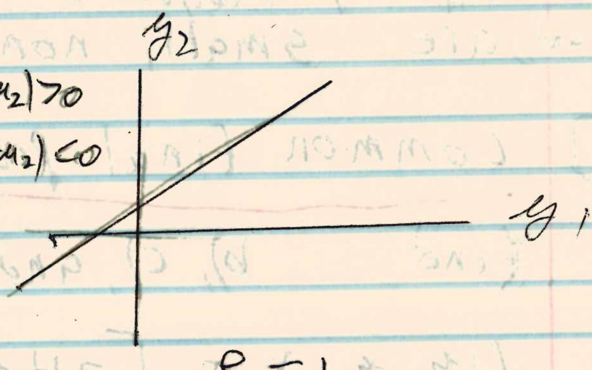
$\rho 3)$ $\rho < 0$ means Y_2 decreases as Y_1 increases.

$\rho 4)$ $\rho = 1$ means all points (Y_1, Y_2) lie on a line with positive slope.

$\rho 5]$ $\rho = -1$ means all points (Y_1, Y_2) lie on a line with negative slope.

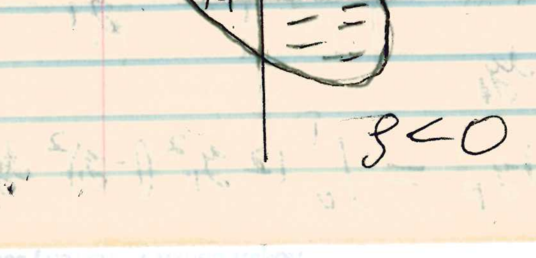


$\rho > 0$

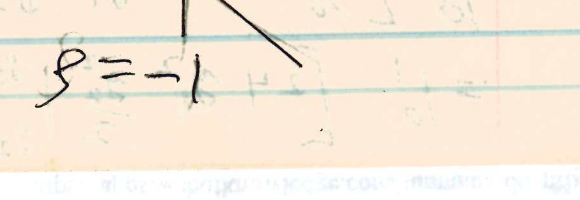


$\rho = 1$

samples from 4 populations of (Y_1, Y_2)



$\rho < 0$



$\rho = -1$

39) Know If Y_1 and Y_2 are independent RVs, then $\operatorname{cov}(Y_1, Y_2) = 0$ if the expected values exist.

proof) $\operatorname{Cov}(Y_1, Y_2) = E(Y_1 Y_2) - E(Y_1)E(Y_2)$

$= E(Y_1)E(Y_2) - E(Y_1)E(Y_2) = 0.$

$\uparrow$ by ind

40) Common Final problem Given

$f(y_1, y_2) = k\, g(y_1, y_2)$ on a triangular or rectangular support,

a) Find k.

b) Find $E(Y_1)$, $E(Y_2)$ and $E(Y_1 Y_2)$.

c) Find $V(Y_1)$ and $V(Y_2)$.

d) Find $\operatorname{cov}(Y_1, Y_2)$.

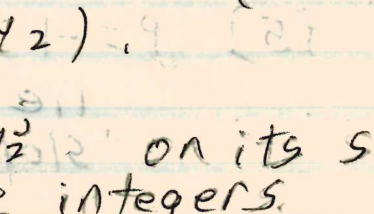
Typically $f(y_1, y_2) = k\, y_1^i\, y_2^j$ on its support where i and j are small nonnegative integers.

41) Common final problem Given $p(y_1, y_2)$

(on a table) find b), c), and d) from point 40).

ex) $f(y_1, y_2) = \begin{cases} 24 y_1 y_2, & 0 \le y_1 \le 1,\ 0 \le y_2 \le 1 \\ & y_1 + y_2 \le 1 \\ 0 & \text{else} \end{cases}$

$E(Y_1) = \int_0^1 \int_0^{1-y_1} y_1\, 24 y_1 y_2 \, dy_2\, dy_1$



$= \int_0^1 \left[\int_0^{1-y_1} 24 y_1^2\, y_2 \, dy_2 \right] dy_1$

$= \int_0^1 \left[24 y_1^2 \frac{y_2^2}{2} \Big|_{y_2=0}^{y_2=1-y_1} \right] dy_1 = \int_0^1 12 y_1^2 (1-y_1)^2 \, dy_1$