$\S 2.8$ 131} p72 Markov's Inequality: If $E(X)$ exists and X has support $\subseteq [0, \infty)$, then for any constant $a > 0$,

$$P(X \geq a) \leq \frac{E(X)}{a}.$$

see proof in book.

132} Chebyshev's inequality: If X is a RV with mean μ and variance σ^2, then for any value $k > 0$, $P(|X-\mu| \geq k) \leq \frac{\sigma^2}{k^2}$.

133} *p73 Strong Law of Large Numbers (SLLN):

Let $X_1, X_2, \ldots$ be iid with $E(X_i) = \mu$. Then

$\frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X} \to \mu$ as $n \to \infty$.

ex} If $k > 0$, $P(|X-\mu| \geq k\sigma) \leq \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}$

So, $P(|X-\mu| < k\sigma) \geq 1 - \frac{1}{k^2}$.

	chebychev	empirical
$P(\lvert X-\mu\rvert < 2\sigma)$	≥ 0.75	$\simeq .95$
$P(\lvert X-\mu\rvert < 3\sigma)$	$\geq \frac{8}{9} \simeq 0.889$	$\simeq .997$

$\underbrace{\hphantom{\simeq .997}}_{\text{if normal}}$

134} Know for final Central Limit Theorem (CLT)!

σ^2

40.5

Let $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$. Then

$$\sqrt{n}(\bar{Y}-\mu) \xrightarrow{D} N(0,\sigma^2),$$

135) *know* Corollary $$\sqrt{n}\left(\frac{\bar{Y}-\mu}{\sigma}\right) = \overbrace{\frac{\bar{Y}-\mu}{\sigma/\sqrt{n}}}^{\text{z score of } \bar{Y}} \xrightarrow{D} N(0,1)$$

and $$\sqrt{n}\left(\frac{\sum_{i=1}^{n} Y_i - n\mu}{n\sigma}\right) = \underbrace{\frac{\sum Y_i - n\mu}{\sqrt{n}\,\sigma}}_{\text{z score of } \sum Y_i} \xrightarrow{D} N(0,1)$$

136) $Y_n \xrightarrow{D} X$ means that for large n, the cdf $F_{Y_n}(y)$ can be approximated by the cdf $F_X(y)$ (at continuity points of $F_X(y)$).

The distribution of X is the limiting distribution of Y_n, and does <u>not</u> depend on n.

137) If $Z_n \xrightarrow{D} N(0,1)$, then the notation $Z_n \approx N(0,1)$, written $Z_n \sim AN(0,1)$, means approximate the cdf of Z_n by the $N(0,1)$ cdf.

$\bar{Y}_n \sim AN(\mu, \frac{\sigma^2}{n})$ means approximate the cdf of $\bar{Y}_n$ as if $\bar{Y}_n \sim N(\mu, \frac{\sigma^2}{n})$. The <u>approximate distribution</u> <u>does depend on n.</u>

... / ... CLT holds forwards and backwards calculations for $\bar{Y} \approx N(\mu, \frac{\sigma^2}{n})$

can be done if i) $Y_1, \ldots, Y_n$ are iid

ii) $E Y_i = \mu$, $V(Y_i) = \sigma^2$

iii) n is large enough

139) Know How large should n be to use the CLT?

i) $n \geq 1$ for $Y_i \sim N(\mu, \sigma^2)$

ii) $n \geq 5$ for Y_i close to (approximately) normal

iii) If Y has a highly skewed distribution (pop), do not use the normal approx if $n \leq 29$.

iv) If $n \geq 100$, usually the CLT holds for this class.

If $30 \leq n \leq 99$ need to be told assume the CLT can be used) or the pop is approx normal.

140) For any n there are distributions where the normal approx is bad, but if n is increased then the normal approx is good. Some researchers recommend $n \geq 5000$ for skewed data.

141) Common E2 problem, Suppose $n = 10$ and $Y_1, \ldots, Y_n$

are iid with $E(Y) = \mu = 100$ and $SD(Y) = \sigma = 15$ $\boxed{4.5}$

a) If Y comes from a highly skewed pop, find

$P(\bar{Y} > 115)$ if possible.

b) If Y comes from a normal dist find

$P(\bar{Y} > 115)$ if possible.

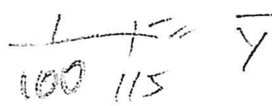
a) not possible, CLT does not apply

b) forwards calculation

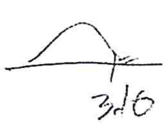
step 0) $\mu_{\bar{Y}} = \mu = 100$, $\sigma_{\bar{Y}} = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{10}} = 4.74$

now $\bar{Y} \approx N(\mu = 100, \sigma_{\bar{Y}} = 4.74)$

do usual forward calculation

i) 

ii) $z = \frac{\bar{Y}\text{value} - \mu}{\sigma/\sqrt{n}} = \frac{115 - 100}{4.74} = 3.16$

iii) 

	.06
3.1	.9992

iv) table $P(\bar{Y} > 115) \approx P(z > 3.16) \approx 1 - P(z < 3.16)$

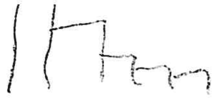
$= 1 - .9992 = \boxed{0.0008}$

common error! $\left\{ \begin{array}{l} \text{use } \sigma \text{ instead of } \sigma/\sqrt{n} \\ \frac{115-100}{15} \text{ instead of } \frac{115-100}{15/\sqrt{n}}. \end{array} \right.$

ex) Y_i iid $EXP(1)$. Get 1000 samples of size n

compute $\bar{Y}_{1n}, \ldots, \bar{Y}_{1000n}$ and make a histogram

to approximate the pdf.

a) $n=1$ 


= pdf of Y for $n=1$

ideal

b) $n=5$


pdf of $\bar{Y}_5$

c) $n=25$


pdf of $\bar{Y}_{25}$

much more normal

See R problem 8 on HW 7.

142] $\bar{Y}$ = mean = sample mean = average = sample average

Ybar

$E(Y)$ = mean = pop mean = average = pop average

Skip §2.9 for now

ch 3 Conditional Probability and Conditional Expectation

§3.2 and §3.3

3] pg 3 * If Y_1 and Y_2 have joint pmf $p(y_1, y_2)$, and marginal pmfs $p_{Y_1}(y_1)$, $p_{Y_2}(y_2)$. Then the condional pmf for Y_1 given $Y_2 = y_2$ is

$$p_{Y_1|Y_2=y_2}(y_1|y_2) = \frac{p(y_1, y_2)}{p_{Y_2}(y_2)} \quad \text{if } p_{Y_2}(y_2) > 0$$

fixed

The conditional pmf for Y_2 given $Y_1 = y_1$ is

$$p_{Y_2|Y_1=y_1}(y_2|y_1) = \frac{p(y_1, y_2)}{p_{Y_1}(y_1)} \quad \text{if } p_{Y_1}(y_1) > 0.$$

fixed

eng Y_1 eng 0	7/26	$\frac{9}{26}$	12/26
noneng 1	5/26	9/26	14/26
$P_{Y_2}(y_2)$	$\frac{8}{26}$	$\frac{18}{26}$	26/26

42.5

$$p(\text{noneng}|F) = p_{Y_1|Y_2=0}(y_1=1|y_2=0) = \frac{p(1,1)}{P_{Y_2}(0)} = \frac{5/26}{8/26} = \frac{5}{8} = 0.675$$

$$= p_{Y_1|Y_2=0}(0|0)$$

$$(F|\text{eng}) = p_{Y_2|Y_1=0}(y_2=0|y_1=0) = \frac{P(0,0)}{P_{Y_1}(0)} = \frac{3/26}{12/26} = \frac{3}{12} = \frac{1}{4} = 0.25$$

$$= p_{Y_2|Y_1=0}(0|0).$$

2) Conditional pmf's are pmfs

fixed

$$\sum_{y_1} p_{Y_1|Y_2=y_2}(y_1|y_2) = 1, \quad \sum_{y_2} p_{Y_2|Y_1=y_1}(y_2|y_1) = 1.$$

3) Know p72 Let Y_1 and Y_2 have joint pdf $f(y_1,y_2)$ and marginal pdfs $f_{Y_1}(y_1)$ and $f_{Y_2}(y_2)$.

Then the conditional pdf of Y_1 given $Y_2 = y_2$ is

$$f_{Y_1|Y_2=y_2}(y_1|y_2) = \frac{f(y_1,y_2)}{f_{Y_2}(y_2)} \quad \text{for any } y_2 \ni f_{Y_2}(y_2) > 0$$

fixed

The conditional pdf of Y_2 given $Y_1 = y_1$ is

$$f_{Y_2|Y_1=y_1}(y_2|y_1) = \frac{f(y_1,y_2)}{f_{Y_1}(y_1)} \quad \text{for any } y_1 \ni f_{Y_1}(y_1) > 0.$$

such that

4) $F(y_1|y_2) = P(Y_1 \le y_1 | Y_2 = y_2)$

$= \int_{-\infty}^{y_1} f_{Y_1|Y_2=y_2}(t_1|y_2)\, dt_1$, $P(a < Y_1 < b | Y_2 = y_2) = \int_a^b f_{Y_1|Y_2=y_2}(y_1|y_2)\, dy_1$

5) Conditional pdfs are pdfs so a RV W can have the same distribution as $Y_1|Y_2=y_2$: $W \sim Y_1|Y_2=y_2$

$\int_{-\infty}^{\infty} f_{Y_1|Y_2=y_2}(y_1|y_2)\, dy_1 = 1$, $\int_{-\infty}^{\infty} f_{Y_2|Y_1=y_1}(y_2|y_1)\, dy_2 = 1$,

6) Common E2 problem: given table for $p(y_1, y_2)$,

i) find $p_{Y_1}(y_1)$, $p_{Y_2}(y_2)$, $p_{Y_1|Y_2=y_2}(y_1|y_2)$, $p_{Y_2|Y_1=y_1}(y_2|y_1)$.

ii) Is $Y_1 \perp\!\!\!\perp Y_2$?

iii) $E(Y_1)$, $E(Y_2)$, $V(Y_1)$, $V(Y_2)$, $E(Y_1Y_2)$, $Cov(Y_1, Y_2)$,

7) Common E2 problem: given $f(y_1, y_2)$

i) find $f_{Y_1}(y_1)$, $f_{Y_2}(y_2)$, $f_{Y_1|Y_2=y_2}(y_1|y_2)$, $f_{Y_2|Y_1=y_1}(y_2|y_1)$

and ii) and iii) of 6).

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 $f_{Y_1}(y_1)$, $f_{Y_2|Y_1=.8}(y_2|0.8)$, and $P(Y_2 \le .5 | Y_1 = .8)$.

soln} Note Y_1 and Y_2 are dependent since $f(y_1, y_2) \ne g(y_1)h(y_2)$.

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

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

435

So $f_{Y_2|Y_1=0.8}(y_2|0.8) = \dfrac{f(0.8, y_2)}{f_{Y_1}(0.8)} = \dfrac{\frac{6}{5}(0.8+y_2^2)}{\frac{6}{5}\left(\frac{1}{3}+0.8\right)} =$

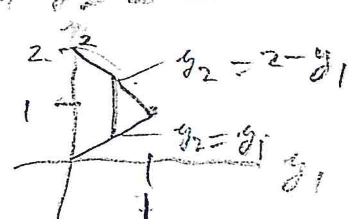
$\dfrac{0.8+y_2^2}{\left(\frac{1}{3}+0.8\right)} = \underbrace{\frac{30}{34}}_{\frac{1}{3}+0.8=\frac{34}{30}}(0.8+y_2^2) = \dfrac{24+30y_2^2}{34}, \quad 0<y_2<1$

$P(Y_2 \le 0.5 \mid Y_1=0.8) = \int_0^{0.5} \frac{24+30y_2^2}{34}\, dy_2$

$= \left(\frac{24}{34} y_2 + \frac{30}{34}\frac{y_2^3}{3}\right)\Big|_0^{0.5} = \frac{12}{34} + \frac{10}{34}\cdot\frac{1}{8} = \underline{0.3897}$

8] With conditional pdfs, the domain can depend on the other variable if the region of integration is not rectangular. The marginal support does not depend on the other variable.

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



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

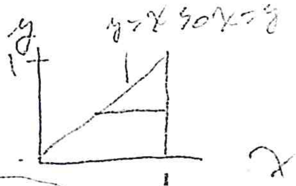
$6y_1^2\,\dfrac{(2-y_1)^2-y_1^2}{2} = 3y_1^2(4-4y_1+y_1^2-y_1^2)$
$= 12y_1^2(1-y_1), \quad 0\le y_1 \le 1.$

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

$$\boxed{\frac{y_2}{2(1-y_1)} \text{ for } \underbrace{y_1 < y_2 < 2-y_1}_{\text{crucial}}} \text{ if } 0 < y_1 < 1,$$

look at the line for integrating

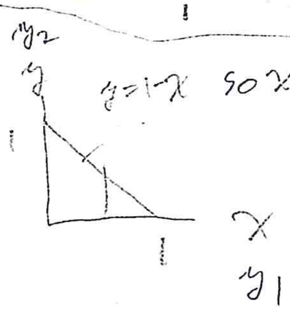
Sec HW 7.



$\int f(x,y)\,dx$
line ∥ x axis

$f_{X|Y}(x|y)$ has support $y<x<1$

$f_{Y|X}(y|x)$ has support $0<y<x$

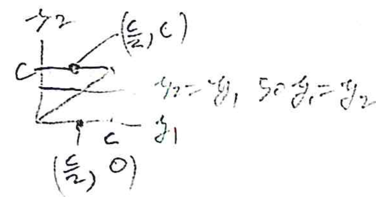


$\int f(x,y)\,dy$
line ∥ y axis

$f_{Y|X}(y|x)$ has support $0<y<1-$

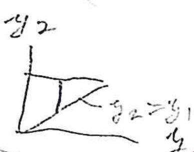
$f_{X|Y}(x|y)$ has support $0<x<1-y$

ex) $f(y_1,y_2)$ :: support $0 \le y_1 \le y_2 \le C$

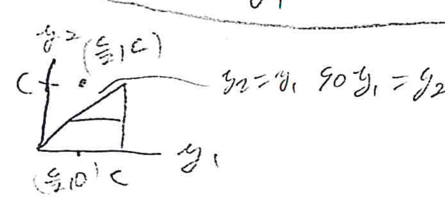


$f_{Y_1|Y_2=y_2}(y_1|y_2)$ has support $0 \le y_1 \le y_2$ (fixed)

$f_{Y_2|Y_1=y_1}(y_2|y_1)$ has support $y_1 \le y_2 \le C$

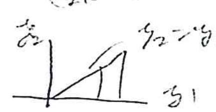


ex) $f(y_1,y_2)$ has support $0 \le y_2 \le y_1 \le C$



$f_{Y_1|Y_2=y_2}(y_1|y_2)$ has support $y_2 \le y_1 \le C$ (fixed)

$f_{Y_2|Y_1=y_1}(y_2|y_1)$ has support $0 \le y_2 \le y_1$



9) $E[g(X) \mid Y=y] = \begin{cases} \int_{-\infty}^{\infty} g(x) f(x \mid y)\, dx & \text{provided } f_Y(y) > 0 \text{ and the integral exists} \\ \sum_x g(x) p(x \mid y) & \text{provided } p_Y(y) > 0 \end{cases}$

$V[X \mid Y=y] = E[X^2 \mid Y=y] - [E(X \mid Y=y)]^2.$

§3.4 10) p100 Let $g(Y) = E(X \mid Y)$ so

$g(y) = E(X \mid Y=y)$ eg wt for people 5' 6' differ.

Let RV $W \sim X \mid Y=y$. Then $X \mid Y$ is a family of RVs.

11) p100 know Iterated Expectations: $E(X) = E[E(X \mid Y)]$.

12) p112 know Conditional Variance Formula

$V(X) = E[V(X \mid Y)] + V[E(X \mid Y)]$.

Note: If $m(y) = E(X \mid Y=y)$ then $m(Y) = E(X \mid Y)$ is a RV
. If $v(y) = V(X \mid Y=y)$ then $v(Y) = V(X \mid Y) = E[X^2 \mid Y] - (E[X \mid Y])^2$ is a RV.

13) p100, 113 $S = S_N = \sum_{i=1}^{N} X_i$ is a compound random variable

if $N \perp\!\!\!\perp X_i$ and the X_i are iid. The dist of N is the compounding dist. The support of $N \subseteq \{0, 1, 2, \dots\}$. If $N=0$, then $S_N = 0$. Then

$E(S_N) = E\left[\sum_{i=1}^{N} X_i\right] = E(N)\, E[X]$ and

$V(S_N) = V\left[\sum_{i=1}^{N} X_i\right] = V(X)\, E(N) + [E(X)]^2\, V(N)$.

proof) $E\left[\sum_{i=1}^{N} X_i\right] = E\left[\underbrace{E\left[\sum_{i=1}^{N} X_i \mid N\right]}\right] = E\left[\sum_{i=1}^{N} \underbrace{E(X_i \mid N)}_{EX_i \text{ since } X_i \perp\!\!\!\perp N}\right]$

given N, N acts like a constant

$= E[N E(X)] \overset{E(X) \text{ is a constant}}{=} E[N]\, E(X)$

Note $E[S_N \mid N] = N\, E[X]$.