§2.8 43]* p51 Mult Law of Prob 4839

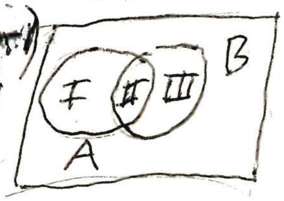
$$P(A \cap B) = P(A)\,P(B|A) = P(B)\,P(A|B).$$

If A and B are ind, $P(A \cap B) = P(A)\,P(B)$.

If $A_1, A_2, \ldots, A_n$ are ind, $P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)\,P(A_2) \cdots P(A_n)$

$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)\,P(A_2|A_1)\,P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$

most general proof use def of cond prob and everything cancels.



44]* Additive Law of Prob.

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

proof: $LHS = P(I) + P(II) + P(III)$

$RHS = P(I) + P(II) + P(II) + P(III) - P(II) = LHS$

45]* If A and B are disjoint, $P(A \cup B) = P(A) + P(B)$ and $P(A \cap B) = 0$

ind, $P(A \cup B) = P(A) + P(B) - P(A)P(B)$

If $A_1, \ldots, A_n$ are pairwise disjoint, $P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + \cdots + P(A_n)$.

46]* complement Rule: $P(A) = 1 - P(\bar{A})$

47] common problem: Given 3 of the 4 probabilities needed in the additive law, find the 4th.

variant: You are told A and B are ind or that A and B are disjoint.

ex] $P(A \cup B) = .8$, $P(A) = .7$, $P(B) = .6$

a) Find $P(A \cap B)$.

b) Are A and B disjoint?

c) Are A and B independent?

soln a) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. so $.8 = .7 + .6 - P(A \cap B)$

or $P(A \cap B) = .7 + .6 - .8 = .5$

b) $P(A) + P(B) = .7 + .6 > 1$ so not disjoint, also $P(A \cap B) \neq 0$

c) $P(A \cap B) = .5 \neq .7(.6)$ so not ind

9.5

ex] $P(A) = .4$, $P(B) = .3$

a) Find $P(A \cup B)$ if A and B are mutually exclusive

b) " " independent.

soln] a) $P(A \cup B) = P(A) + P(B) = .4 + .3 = .7$

b) $P(A \cup B) = P(A) + P(B) - P(A \cap B) =$

$.4 + .3 - (.4)(.3) = .58$

§2.9 48] p62, 64 The <u>event composition</u> method for finding $P(A)$ expresses the event A as a composition of 2 or more events using unions, intersections and complements. Then the additive and multiplicative laws of prob are applied to find $P(A)$.

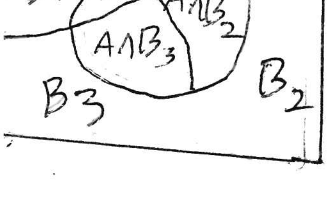
§ 2.10

ex] p71 Let $S = B_1 \cup B_2 \cup \dots \cup B_K$ and $P(B_i) > 0$ for $i = 1, \dots, K$.

where $B_i \cap B_j = \phi$ for $i \neq j$

Let $A \subseteq S$. Then $A = \underbrace{(A \cap B_1) \cup (A \cap B_2) \cup \dots \cup (A \cap B_K)}_{\text{disjoint}}$

Thus $P(A) = \sum_{i=1}^{K} \underbrace{P(A \cap B_i)}_{= P(A|B_i)P(B_i)} = \sum_{i=1}^{K} P(A|B_i)\, P(B_i)$.



This formula is called the <u>law of total probability</u> and is useful when $P(A|B_i)$ and $P(B_i)$ are given for $i = 1, \dots, K$. See ex 2.17.

<u>Bayes rule</u> $P(B_j|A) = \dfrac{P(A \cap B_j)}{P(A)} = \dfrac{P(A|B_j)\, P(B_j)}{\sum_{i=1}^{K} P(A|B_i)\, P(B_i)}$

is a corollary.

ex] Perform n ind. identical experiments.

Interest is in a single outcome of the experiment, say D. better way soon b.n RV

i) Find P(D occurred in none of the n expts).

ii) Find P (D occurred in at least one of the n expts)

iii) Find P (D occurred in all n expts)

iv) Find P (D occurred in not all n expts)

Soln i) $P(\text{none}) = \underbrace{\overbrace{1-P(D)}^{\text{complement rule}}}_{1} \; \underbrace{1-P(D)}_{2} \cdots \underbrace{1-P(D)}_{n} = [1-P(D)]^n$

$= P(\overline{D_1} \wedge \overline{D_2} \wedge \cdots \wedge \overline{D_n})$ where D_i means D occurred in ith expt

ii) $P(\text{at least one}) = P(\text{not none}) = 1 - P(\text{none}) = 1 - [1-P(D)]^n$

$= P(\overline{\overline{D_1} \wedge \overline{D_2} \wedge \cdots \wedge \overline{D_n}})$

iii $P(\text{all}) = P(D_1 \wedge D_2 \wedge \cdots \wedge D_n) = [P(D)]^n$

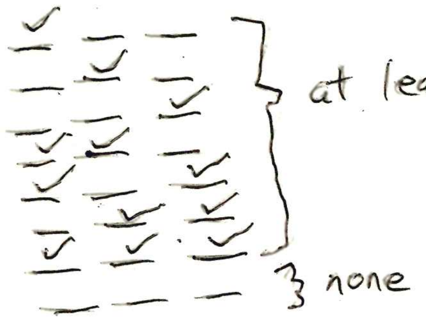
iv) $P(\text{not all}) = 1 - P(\text{all}) = 1 - [P(D)]^n = P(\overline{D_1 \wedge D_2 \wedge \cdots \wedge D_n})$

see ex. 2.20 on p. 65-6 and ex 2.22a

ex] roll die 3 times D = die was a 5

i) P(none of the rolls are 5's) $= \frac{5}{6}\frac{5}{6}\frac{5}{6} = (1-\frac{1}{6})^3 = (\frac{5}{6})^3$

ii) P(at least one of the rolls is a 5) $= 1 - (\frac{5}{6})^3$



✓ ⟺ D_i occurred

blank ⟺ D_i did not occur $= \overline{D_i}$

P(at least one) $= 1 - P$(none)

iii) P(all 3 rolls are 5's) $= \underset{1st}{\frac{1}{6}} \; \underset{2nd}{\frac{1}{6}} \; \underset{3rd}{\frac{1}{6}} = (\frac{1}{6})^3$

iv) P(not all 3 rolls are 5's) $= 1 - (\frac{1}{6})^3$

ex] 70 rolls P(none are 5's) $= (\frac{5}{6})^{70}$

Skip ex. 2.19 and 2.21. 10.7

49] p76, 86* A <u>random variable</u> is a real valued function for which the domain is a sample space.

50] p77-78 The entire group of objects that we want information about is the <u>population</u>. A <u>sample</u> is the part of the population we actually examine in order to gather information. Let N and n denote the pop and sample sizes. If each of the $\binom{N}{n}$ samples is equally likely, the result is a <u>random sample</u>.

ex] A <u>voluntary response sample</u> consists of people who choose themselves by responding to a general appeal. People with strong opinions are more likely to respond. Ann Landers asked her readers if they had it to do over again, would they have kids. Almost 10000 parents responded and 70% said no. In a random sample, about 9% would say no.

say in words

ex] Flip fair coin twice $S = \{HH, HT, TH, TT\}$

Let Y = # of heads $Y(HH) = 2$, $Y(HT) = 1$, $Y(TH) = 1$ $Y(TT) = 0$

y	0	1	2
$P(Y = y)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Ch 3 1] p87* A RV is <u>discrete</u> if it can assume only a finite or countably infinite number of distinct values. The collection of these probabilities is the <u>probability distribution</u> of the discrete RV.

2] Notation uppercase Y, X for RV lowercase for the values the RV may assume.

3) p88 The event $\{Y = y\} = \{E_i \in S : Y(E_i) = y\}$.

Hence $P(Y = y)$ is the sum of the prob's of all of the sample points in S that are assigned the value y. Let $p(y) = P(Y = y)$ be the probability function of discrete Y.

see ex where coin is tossed twice, Y = # heads.

4) The probability distribution for a discrete RV Y can be given by a table or formula that gives $P(Y = y) = p(y)$ for all y. Generally, if $p(y) = 0$, that value of y is omitted.

5) p89 * For the collection of $p(y)$ to form a discrete prob. dist., the following must hold:

1) $0 \le p(y) \le 1 \quad \forall y$

2) $\sum_y p(y) = 1$ where the sum is over all values of $p(y) > 0$.

6] Common problem give table with a $p(y)$ omitted.

ex]

y	0	1	2	3
$p(y)$	.1	.2	.3	

then $p(3) = .4 = 1 - .1 - .2 - .3$.

7] A probability distribution is a model or hopefully useful approximation for the data population.

3.3 8] Know Let Y be a discrete RV with probability function $p(y)$. Then the expected value of Y is $E(Y) = \sum_{\underbrace{y}_{y:\, p(y) > 0}} y \, p(y)$.

9) Common problem: given table of y and $p(y)$, find $E(Y)$. 115

ex]

y	0	1
$p(y)$	$\frac{1}{2}$	$\frac{1}{2}$

$E(Y) = \Sigma y p(y) = 0(\frac{1}{2}) + 1(\frac{1}{2}) = \frac{1}{2}$

ex]

y	-1	1
$p(y)$	$\frac{1}{2}$	$\frac{1}{2}$

$E(Y) = -1(\frac{1}{2}) + 1(\frac{1}{2}) = 0$

ex]

y	-2	-1
$p(y)$	$\frac{1}{2}$	$\frac{1}{2}$

$E(Y) = -2(\frac{1}{2}) + -1(\frac{1}{2}) = -\frac{3}{2}$

10] $E(Y) = \mu$ = population mean is a measure of the location of the population.

ex] mean ht of adult US women
men
newborns 0 age 20 in 8 lbs
wt
age
networth
yearly income

11] p93 Know If Y is a discrete valued RV with $p(y)$ and if $g(Y)$ is a real valued function of Y, then $g(Y)$ is a RV and $E\,g(Y) = \sum_{y:p(y)>0} g(y)\,p(y)$

skim proof

12] p93 Know The variance of a RV Y is $V(Y) = E\,(Y - EY)^2$. The standard deviation (SD) of Y is the positive square root of $V(Y)$.

13] $V(Y) = \sigma^2$, $SD(Y) = \sigma = \sqrt{V(Y)}$ = SD of Y.

$V(Y) = \Sigma\,(y - E(Y))^2 p(y)$ so $g(y) = (y - E(Y))^2$

14] $\boxed{\text{Common final Problem}}$ given a table of y and $p(y)$, find $E(g(Y))$ and σ.

ex]

y	7
$P(y)$	1

$y-E(Y)$	0
$P(y-E(Y))$	1

$E(Y) = 7 \cdot 1 = 7$ $\qquad Var(Y) = 0$ $\qquad SD(Y) = \sqrt{0} = 0$

ex] Quiz 2 may ask you to make a table to compute σ

y	$P(y)$	$y \cdot P(y)$	$y-E(Y)$	$(Y-E(Y))^2\, P(y)$
3	.3	.9	−1	.3
4	.4	1.6	0	0
5	.3	1.5	1	.3
sums: $\Sigma P(y)=1$		$E(Y)=4$		$.6 = V(Y) = \sum_{y:P(y)>0} (y-E(Y))^2 P(y)$

$SD(Y) = \sqrt{.6} = .7746$

$\Sigma\, y-E(Y)$ is meaningless but $\Sigma (y-EY)P(y) = E(Y-EY) = 0$

ex]

y	$P(y)$	$y \cdot P(y)$	$y-E(Y)$	$(Y-E(Y))^2\, P(y)$
1	.4	.4	−3	3.6
2	.1	.2	−2	.4
6	.3	1.8	2	1.2
8	.2	1.6	4	3.2
sums:	$\Sigma P(y)=1$	$E(Y)=4.0$		$8.4 = V(Y) = \sum_{y:P(y)>0} (y-E(Y))^2 P(y)$

$E(Y) = 4.0 \qquad \sigma = \sqrt{8.4} = 2.8983$

After making a table or two, you may just say $\sigma^2 = \Sigma (y-EY)^2 P(y) = (1-4)^2(.4) + (2-4)^2(.1) + (6-4)^2(.3) + (8-4)^2(.2) = 8.4$

15] σ is a measure of the spread of the population.

16] Empirical rule p10 (68-95-99.7 or) For many bell shaped populations,

$\mu \pm \sigma$ contains $\approx$ 68% of measurements

$\mu \pm 2\sigma$ $\approx$ 95%

$\mu \pm 3\sigma$ $\approx$ 99.7% $\approx$ all

17] *Expectation rules 125

$E(c) = c$ if $g(Y) \equiv c$ is a constant

$E[c\,g(Y)] = c\,E(g(Y))$

$E\left[\sum_{i=1}^{k} g_i(Y)\right] = \sum_{i=1}^{k} E[g_i(Y)]$.

18] p96 Know short cut formula for variance

$$V(Y) = E[Y - E(Y)]^2 = E(Y^2) - [E(Y)]^2$$

proof: $E[(Y-\mu)^2] = E[Y^2 - 2Y\mu + \mu^2]$

$= E(Y^2) - 2\mu E(Y) + \mu^2 = EY^2 - 2\mu^2 + \mu^2$

$= E(Y^2) - \mu^2$

19] Know $E[Y^2] = V(Y) + [E(Y)]^2$.

ex) Table

y	y^2	$p(y)$	$y \cdot p(y)$	$y^2\,p(y)$
1	1	.4	.4	.4
2	4	.1	.2	.4
6	36	.3	1.8	10.8
8	64	.2	1.6	12.8
			$E(Y) = 4.0$	$24.4 = E(Y^2)$

$V(Y) = E(Y^2) - (E(Y))^2 = 24.4 - 16 = 8.4$

so $\sigma = \sqrt{8.4}$ as before

Also, $E(Y^2) = (1)^2(.4) + (2)^2(.1) + (6)^2(.3) + 8^2(.2) = 24.4$

section §3.4 20] p101 An expt is a binomial experiment if

1) the expt consists of n identical trials
2) each trial has 2 outcomes "success"s and F
3) For each trial, $P(\text{Success}) = p$ ← no longer sample space, $P(F) = 1-p$
4) The trials are independent
5) RV Y = number of successes in the n trials

Note: "success" does not have the dictionary meaning. It is what you count eg males, females, number of people that die after being diagnosed with colon cancer.

ex] Flip coin, count # of heads

ex] polls with random samples of size $n = 100$
Y counts # for Clinton. This expt is approx binomial (the trials are approximately ind if the sample size $n << N =$ pop size)

ex] similarly randomly select 1000 people, Y counts the number that test HIV positive

ex] Ann Landers voluntary response sample is not a binomial expt (no pop is represented, no ind)

ex] Prof needs students to comment on his teaching and picks several A students. $Y = \#$ that give high rating. This is not a bin expt. (pop of prof students is not represented, no ind)

Verbally

21] p103 Know RV Y has a binomial distribution with success probability p, based on n trials written $Y \sim \text{bin}(n,p)$, if $f(y) = \binom{n}{y} p^y (1-p)^{n-y}$

for $y = 0, 1, \ldots, n$.

Note: $q = 1-p$. If there are n trials and y S's, then there are $n-y$ F's.

Prob (y S's, $n-y$ F's) $= p^y (1-p)^{n-y}$ by ind

There are $\binom{n}{y}$ sequences with y S's.

By the binomial theorem, $1 = (p + 1-p)^n = \sum_{y=0}^{n} \binom{n}{y} p^y (1-p)^{n-y}$.

so $f(y)$ is a prob. fn. (see p104 & p.46)

22] p 107 Know If $Y \sim bin(n,p)$, then 13.5

$E(Y) = np$ and $V(Y) = np(1-p)$.

Proof] $$EY = \sum_{y=0}^{n} y \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y}$$

$$= np \sum_{y=1}^{n} \frac{(n-1)!}{(y-1)!\,(n-y)!} p^{y-1} (1-p)^{n-y}.$$

Let $z = y-1$ then $\left.\begin{array}{l} y=n \to z=n-1 \\ y=1 \to z=0 \end{array}\right\}$, like u substitution

and $$E(Y) \overset{\text{trick}}{=} np \sum_{z=0}^{n-1} \frac{(n-1)!}{z!\,(n-1-z)!} p^z (1-p)^{n-1-z}$$

$$= np \underbrace{\sum_{z=0}^{n-1} \binom{n-1}{z} p^z (1-p)^{n-1-z}}_{1 = \sum_z p_{bin(n-1,p)}(z)} \overset{\text{trick}}{=} np$$

Now $EY^2 = E[Y(Y-1)] + \mu$ and

$$EY(Y-1) \overset{\text{trick}}{=} \sum_{y=0}^{n} y(y-1) \frac{n!}{y!\,(n-y)!} p^y q^{n-y}$$

$$\overset{\text{trick}}{=} n(n-1)p^2 \sum_{y=2}^{n} \frac{(n-2)!}{(y-2)!\,(n-y)!} p^{y-2} q^{n-y}$$

Let $z = y-2$. $\begin{array}{l} y=n \to z=n-2 \\ y=2 \to z=0 \end{array}$

$$E(Y(Y-1)) = n(n-1)p^2 \underbrace{\sum_{z=0}^{n-2} \frac{(n-2)!}{z!\,(n-2-z)!} p^z q^{n-2-z}}_{1 = \sum p_{bin(n-2,p)}(y)}$$

So $EY(Y-1) \overset{\text{trick}}{=} n(n-1)p^2$ and $EY^2 = n(n-1)p^2 + np$

$V(Y) = EY^2 - (EY)^2 = n(n-1)p^2 + np - n^2p^2 = \ldots = np(1-p)$.