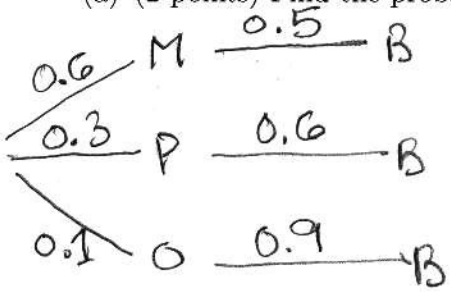
1. Of the travelers arriving at a small airport, 60% fly on major airlines, 30% fly on privately owned planes, and the remainder fly on commercially owned planes not belonging to a major airline. Of those traveling on major airlines, 50% are traveling for business reasons, whereas 60% of those arriving on private planes and 90% of those arriving on other commerially owned planes are traveling for business reasons. Suppose we select one person arriving at this airport.

   (a) (2 points) Find the probability that the person is traveling on business.



$$P[B] = \left(\frac{6}{10}\right)\left(\frac{5}{10}\right) + \left(\frac{3}{10}\right)\left(\frac{6}{10}\right) + \left(\frac{1}{10}\right)\left(\frac{9}{10}\right)$$

$$= \frac{30+18+9}{100} = \boxed{\frac{57}{100}}$$

   (b) (2 points) Find the probability that the person is traveling for business on a privately owned plane.

$$P[B,P] = \left(\frac{3}{10}\right)\left(\frac{6}{10}\right) = \frac{18}{100} = \boxed{\frac{9}{50}}$$

   (c) (3 points) Find the probability that the person arrived on a privately owned plane, given that the person is flying on a commercially owned plane.

$$P[P|C] = \boxed{0}$$

   (d) (3 points) Given that the person is traveling on business, what is the probability that the flight is on a private plane?

$$P[P|B] = \frac{P[B,P]}{P[B]} = \boxed{\frac{18}{57} \approx 0.3158}$$

2. Ten percent of the engines manufactured on an assembly line are defective. Assuming that the number of engines manufactured is extremely large, answer the following.

(a) (2 points) If engines are randomly selected one at a time and tested, what is the probability that the first working engine will be found on the second trial?

N = # of engines tested until first working engine

N is geometric $(0.9) \Rightarrow P_N(2) = \left(\frac{1}{10}\right)\left(\frac{9}{10}\right) = \boxed{\frac{9}{100}}$

(b) (3 points) What is the probability that the third working engine will be found on or before the fifth trial?

M = # of engines tested until 3rd working engine

M is Pascal $(3, 0.9) \Rightarrow P[M \le 5] = P_M(3) + P_M(4) + P_M(5)$

$$P[M \le 5] = \left(\frac{9}{10}\right)^3 + \binom{3}{2}\left(\frac{1}{10}\right)\left(\frac{9}{10}\right)^3 + \binom{4}{2}\left(\frac{1}{10}\right)^2\left(\frac{9}{10}\right)^3 \approx \boxed{0.9914}$$

(c) (2 points) Find the mean and variance of the number of trials on which the third working engine is found.

$$E[M] = \frac{3}{9/10} = \boxed{\frac{30}{9} = 3.3\overline{3}}$$

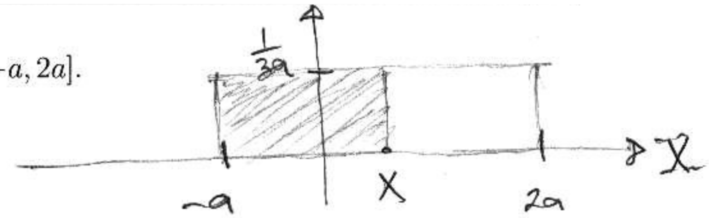
$$Var[M] = \frac{3\left(\frac{1}{10}\right)}{(9/10)^2} = \frac{3}{10} \cdot \frac{100}{81} = \boxed{\frac{10}{27} \approx 0.\overline{370}}$$

(d) (3 points) Given that the first two engines tested were defective, what is the probability that at least two more engines must be tested before the first working engine is found?

Trials are INDEPENDENT b/c # of engines manufactured is very large, so same as

$$P[N \ge 2] = 1 - P[N < 2] = 1 - P_N(1) = 1 - \frac{9}{10} = \boxed{\frac{1}{10}}$$

3. Let X be uniformly distributed over the interval $[-a, 2a]$.



(a) (4 points) Derive the CDF of X.

If $-a \le x \le 2a$, then

$$F_X(x) = P[X \le x] = \frac{1}{3a}\left[x - (-a)\right] = \frac{x+a}{3a}$$

$$F_X(x) = \begin{cases} 0 & x < -a \\ \frac{x+a}{3a} & -a \le x \le 2a \\ 1 & x > 2a \end{cases}$$

(b) (3 points) Let $Y = \dfrac{X}{3} + 1$. Find $E[Y]$ and $Var[Y]$.

$$E[Y] = \frac{1}{3}E[X] + 1 = \frac{1}{3}\left(\frac{a}{2}\right) + 1 = \boxed{\frac{a}{6} + 1}$$

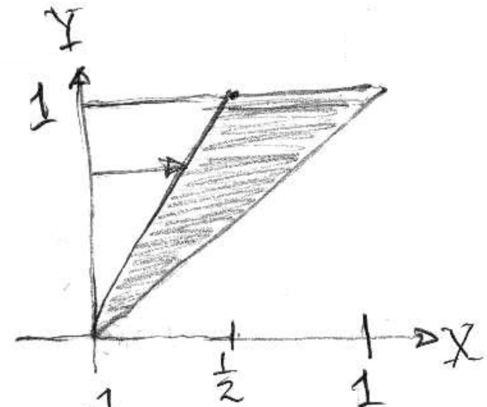
$$Var[Y] = \left(\frac{1}{3}\right)^2 Var[X] = \frac{1}{9} \cdot \frac{(3a)^2}{12} = \boxed{\frac{a^2}{12}}$$

(c) (3 points) Find $P\left[-\dfrac{a}{3} + 1 \le Y \le 1\right]$.

$$P\left[-\frac{1}{3}a + 1 \le \frac{1}{3}X + 1 \le 1\right] = P\left[-\frac{1}{3}a \le \frac{1}{3}X \le 0\right]$$

$$= P[-a \le X \le 0] = \boxed{\frac{1}{3}}$$

4. Random variables X and Y have joint PDF



$$f_{X,Y}(x,y) = \begin{cases} c(x+y) & 0 \le x \le y \le 1 \\ 0 & otherwise \end{cases}$$

(a) (3 points) Find the value of the constant c.

$$\int_0^1 \int_0^y c(x+y)\,dx\,dy = c\int_0^1 \left(\frac{x^2}{2}+xy\right)\Big|_0^y dy = c\int_0^1 \frac{3}{2}y^2\,dy$$

$$= \frac{c}{2}\cdot y^3\Big|_0^1 = \frac{c}{2} = 1 \Rightarrow \boxed{c=2}$$

(b) (4 points) Find $P\left[Y \le 2X\right] = 1 - P[Y > 2X]$

$$= 1 - \int_0^1 \int_0^{y/2} 2(x+y)\,dx\,dy = 1 - \int_0^1 (x^2+2xy)\Big|_0^{y/2} dy$$

$$= 1 - \int_0^1 \left(\frac{1}{4}y^2 + y^2\right)dy = 1 - \frac{5}{4}\cdot\frac{y^3}{3}\Big|_0^1$$

$$= 1 - \frac{5}{12}(1-0) = \boxed{\frac{7}{12}}$$

(c) (3 points) Find $f_Y(y)$.

$\underline{0 \le y \le 1}$: $f_Y(y) = \int_0^y 2(x+y)\,dx = x^2+2xy\Big|_0^y = 3y^2$

$$f_Y(y) = \begin{cases} 3y^2 & 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

5. In response to requests for information, a company sends faxes that can be 1, 2 or 3 pages in length depending on the information requested. The PMF of L, the length of one fax is

$$P_L(l) = \begin{cases} 1/3 & l=1 \\ 1/2 & l=2 \\ 1/6 & l=3 \\ 0 & otherwise \end{cases}$$

For a set of four independent information requests:

(a) (4 points) Find the joint PMF of the random variables X, Y and Z, the number of 1-page, 2-page and 3-page faxes, respectively.

$$P_{X,Y,Z}(x,y,z) = \binom{4}{x,y,z} \left(\frac{1}{3}\right)^x \left(\frac{1}{2}\right)^y \left(\frac{1}{6}\right)^z$$

(b) (3 points) In a group of four faxes, what is the PMF of the number of 3-page faxes?

Z is binomial $(4, \frac{1}{6})$:

$$P_Z(z) = \binom{4}{z} \left(\frac{1}{6}\right)^z \left(\frac{5}{6}\right)^{4-z}$$

(c) (3 points) Given that there are two 3-page faxes in a group of four, what is the expected number of 1-page faxes?

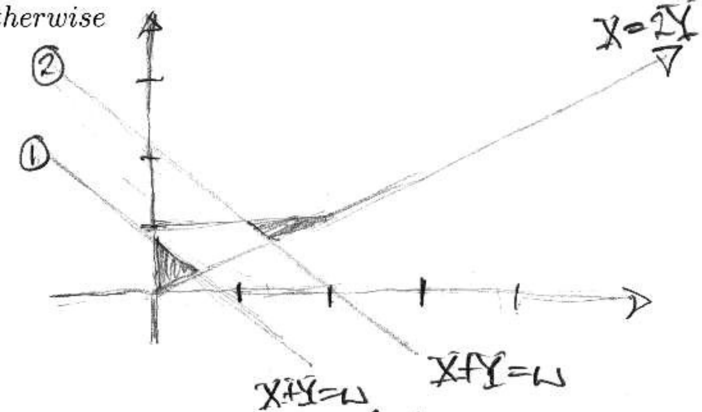
$$P_{X|Z}(x|2) = \binom{2}{x} \left(\frac{2}{5}\right)^x \left(\frac{3}{5}\right)^{2-x}$$

That is, given $Z=2$, X is binomial $(2, \frac{2}{3})$

$$\Rightarrow E[X|Z=2] = 2 \cdot \frac{2}{5} = \boxed{\frac{4}{5}}$$

6. Let X and Y have joint PDF

$$f_{X,Y}(x,y) = \begin{cases} 1 & 0 \le x \le 2,\ 0 \le y \le 1,\ x \le 2y \\ 0 & otherwise \end{cases}$$



Let $W = X + Y$.

(a) (2 points) Find the range of W.

$S_W = [0, 3]$

(b) (6 points) Find the CDF of W.

① $0 \le w \le 1$:

$$F_W(w) = \int_0^{2w/3} \int_{x/2}^{w-x} dy\,dx = \int_0^{2w/3} \left(w - \frac{3x}{2}\right) dx = wx - \frac{3}{4}x^2 \Big|_0^{2w/3} = \left(\frac{2}{3} - \frac{1}{3}\right) w^2 = \frac{1}{3} w^2$$

② $1 \le w \le 3$:

$$F_W(w) = 1 - \int_{w/3}^{1} \int_{w-y}^{2y} dx\,dy = 1 - \int_{w/3}^{1} (3y - w)\,dy = 1 + \left(wy - \frac{3}{2}y^2\right)\Big|_{w/3}^{1}$$

$$= 1 + w - \frac{1}{3}w^2 - \frac{3}{2} + \frac{1}{6}w^2 = w - \frac{1}{2} - \frac{1}{6}w^2$$

$$\Longrightarrow F_W(w) = \begin{cases} 0 & w < 0 \\ \frac{1}{3}w^2 & 0 \le w \le 1 \\ w - \frac{1}{2} - \frac{1}{6}w^2 & 1 \le w \le 3 \\ 1 & \text{otherwise} \end{cases}$$

(c) (2 points) Find the PDF of W.

$$f_W(w) = \begin{cases} \frac{2}{3}w & 0 < w < 1 \\ 1 - \frac{1}{3}w & 1 < w < 3 \\ 0 & \text{otherwise} \end{cases}$$

7. Let X_A be the indicator random variable for event A with probability $P[A] = 0.7$. Let $\hat{P}_n(A)$ denote the relative frequencey of event A in n independent trials.

(a) (3 points) Find $E[X_A]$ and $Var[X_A]$.

X_A is Bernoulli $(0.7) \Rightarrow E[X_A] = \boxed{\frac{7}{10}}$

and $Var[X_A] = \frac{7}{10} \cdot \frac{3}{10} = \boxed{\frac{21}{100}}$

(b) (3 points) Find $Var[\hat{P}_n(A)]$.

$Var[\hat{P}_n(A)] = Var[M_n(X_A)] = \frac{1}{n} Var[X_A] = \boxed{\frac{21}{100n}}$

(c) (4 points) Use the Chebyshev inequality to find the confidence coefficient $1 - \alpha$ such that $\hat{P}_{100}(A)$ is within 0.1 of $P[A]$. In other words, find α such that

$$P[|\hat{P}_{100}(A) - P[A]| \leq 0.1] \geq 1 - \alpha.$$

$P[|\hat{P}_n(A) - P[A]| \geq c] \leq \frac{Var[\hat{P}_n(A)]}{c^2} = \frac{21}{100n \cdot c^2} = \alpha$

$n = 100$ and $c = \frac{1}{10} \Rightarrow n \cdot c^2 = 1 \Rightarrow \alpha = \frac{21}{100}$

$\Rightarrow 1 - \alpha = \boxed{\frac{79}{100}}$

8. (5 points) Let K be the number of heads in $n = 200$ flips of a coin. Devise a significance test for the hypothesis H that the coin is fair such that the signifcance level is $\alpha = 0.05$ and the rejection set R has the form $\{|K - E[K]| > c\}$. Use the Central Limit Theorem.

H_0: coin is fair; i.e., K is Bernoulli$(200, \frac{1}{2}) \Rightarrow E[K] = 100$.

WANT: $P[|K-\mu_K| \leq c] \geq 0.95$ and $\mathrm{Var}[K] = 50$

$$P[|K-\mu_K| \leq c] = P\left[-\frac{c}{\sigma_K} \leq \frac{K-\mu_K}{\sigma_K} \leq \frac{c}{\sigma_K}\right] \approx P\left[-\frac{c}{5\sqrt{2}} \leq Z \leq \frac{c}{5\sqrt{2}}\right]$$

and $P\left[-\frac{c}{5\sqrt{2}} \leq Z \leq \frac{c}{5\sqrt{2}}\right] = 2\Phi\left(\frac{c}{5\sqrt{2}}\right) - 1 \geq 0.95 \Rightarrow \Phi\left(\frac{c}{5\sqrt{2}}\right) \geq 0.975$

$\Rightarrow \frac{c}{5\sqrt{2}} \geq 1.96 \Rightarrow c \geq 13.9 \Rightarrow c = 14$

If $86 \leq K \leq 114$, the coin is FAIR with 95% probability.

9. (5 points) The duration of a data telephone call is an exponential random variable D with $E[D] = 5$ minutes. The duration of a voice call is an exponential random variable V with $E[V] = \mu_V > 5$ minutes. The null hypothesis of a binary hypothesis test is H_0: a call is a data call. The alternative hypothesis is H_1: a call is a voice call. The probability of a data call is $P[D] = 0.65$ and the probability of a voice call is $P[V] = 0.35$. A binary hypothesis test measures T minutes, the duration of a call. The decision is H_0 if $T \leq t_0$ minutes. Otherwise, the decision is H_1. Design a MAP hypothesis test for when $\mu_V = 7$ minutes. Clearly state your decision rules in terms of data and voice calls.

H_0: data call $\Rightarrow T$ is exponential$(1/5)$

H_1: voice call $\Rightarrow T$ is exponential$(1/7)$

$$\frac{f_{T|H_0}(t)}{f_{T|H_1}(t)} = \frac{\frac{1}{5}e^{-t/5}}{\frac{1}{7}e^{-t/7}} = \frac{7}{5}e^{t/7 - t/5} \geq \frac{P[H_1]}{P[H_0]} = \frac{0.35}{0.65} = \frac{7}{13}$$

$$-\frac{2t}{35} \geq \ln\left(\frac{5}{13}\right) \Rightarrow t \leq -\frac{35}{2}\ln(5/13) \approx 16.7$$

If $t \leq 16.7$, conclude that the call is DATA; otherwise, conclude that the call is VOICE.

10. The following table gives $P_{X,Y}(x, y)$, the joint probability mass function of the random variables X and Y.

$P_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$	
$x=-1$	1/10	0	1/10	0.2
$x=0$	0	4/10	0	0.4
$x=1$	2/10	0	2/10	0.4
	0.3	0.4	0.3	

(a) (5 points) Find $Var[X]$, $Var[Y]$, and $Cov[X, Y]$.

$$E[X] = (-1)\left(\tfrac{2}{10}\right) + 0 + \tfrac{4}{10} = \tfrac{1}{5} \qquad E[X^2] = (-1)^2\left(\tfrac{2}{10}\right) + 0 + \left(\tfrac{4}{10}\right) = \tfrac{3}{5}$$

$$E[Y] = 0 + \tfrac{4}{10} + (2)\left(\tfrac{3}{10}\right) = 1 \qquad E[Y^2] = 0 + \tfrac{4}{10} + (2)^2\left(\tfrac{3}{10}\right) = \tfrac{8}{5}$$

$$Var[X] = E[X^2] - (\mu_X)^2 = \tfrac{3}{5} - \tfrac{1}{25} = \boxed{\tfrac{14}{25}}$$

$$Var[Y] = E[Y^2] - (\mu_Y)^2 = \tfrac{8}{5} - 1 = \boxed{\tfrac{3}{5}}$$

$$E[XY] = (-1)(2)\left(\tfrac{1}{10}\right) + (1)(2)\left(\tfrac{2}{10}\right) = \tfrac{1}{5}$$

$$Cov[X,Y] = E[XY] - \mu_X \mu_Y = \tfrac{1}{5} - \tfrac{1}{5} = \boxed{0}$$

(b) (5 points) Find a^* and b^*, the optimum coefficients in the linear estimate $\hat{X}_L(Y) = a^*Y + b^*$.

$$Cov[X,Y] = 0 \Rightarrow \hat{X}_L(Y) = \hat{X}_B = E[X] = \boxed{\tfrac{1}{5}}$$