Logistical and Transportation Planning Methods

Solution
Problem Set #1

Problem 1
Two-horse race

(a). The conditional pdf of U given that $V = v$ is:

$$f_{U|V}(u|v) = \frac{f(u,v)}{f_V(v)}$$ Equation 1

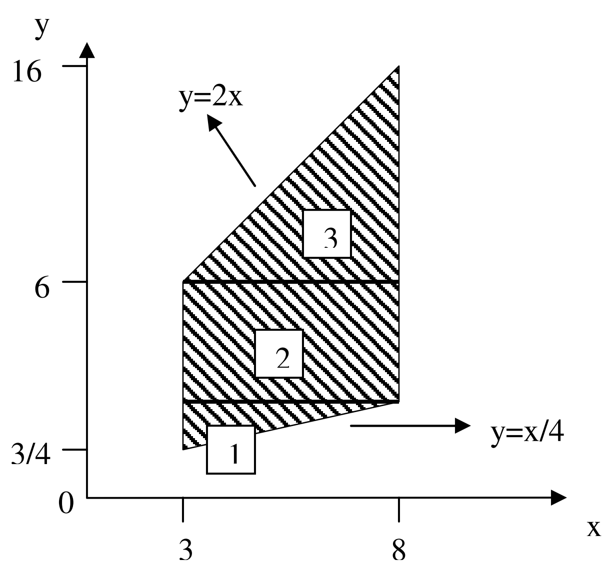
The marginal pdf of V is given by:

$$f_V(v) = \int_{-\infty}^{\infty} f(u,v)du$$ Equation 2

(b). There are two random variables:

- X , which is the finishing time for A, with $x \in [3,8]$ in minutes;
- Y , which is the finishing time for B, with $y \in \left[\frac{x}{4};2x\right]$ in minutes.

The joint sample space is therefore:



The probability law of the sample space is:

$$f_{X,Y}(x,y) = f_{Y|X}(y|x) \cdot f_X(x) = \frac{4}{7x} \cdot \frac{1}{5} = \frac{4}{35x}$$ Equation 1

Then from question a):

$$\forall y \in \left[\frac{3}{4};16\right] \qquad f_Y(y) = \int_3^8 f_{X,Y}(x,y)dx$$ Equation 2

From the joint sample space, we can see that there are 3 different cases for the boundaries of the integration:

Case #1: $\forall y \in \left[\frac{3}{4};2\right]$ $\quad f_Y(y) = \int_3^{4y} f_{X,Y}(x,y)dx = \int_3^{4y} \frac{4}{35x}dx = \frac{4}{35}\ln\frac{4y}{3}$

Case #2 $\forall y \in [2;6]$ $\quad f_Y(y) = \int_6^{8} f_{X,Y}(x,y)dx = \frac{4}{35}\ln\frac{8}{3}$

Case #3 $\forall y \in [6;1\ \ ]$ $\quad f_Y(y) = \int_{y/2}^{8} f_{X,Y}(x,y)dx = \frac{4}{35}\ln\frac{16}{y}$

(c).

$$f_{X|Y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y} \qquad \forall y \in \left[\frac{3}{4};16\right] \qquad \text{Equation 1}$$

There are three different cases:

Case #1: $\forall y \in \left[\frac{3}{4};2\right]$ $\quad f_{X|Y}(x,y) = \frac{4/35x}{\frac{4}{35}\ln\frac{4y}{3}} = \frac{1}{x \cdot \ln\frac{4y}{3}}$

Case #2 $\forall y \in [2;6]$ $\quad f_{X|Y}(x,y) = \frac{1}{x \cdot \ln\frac{8}{3}}$

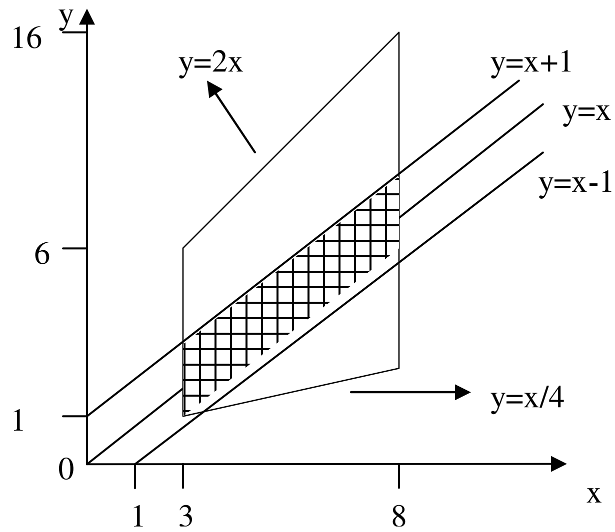
Case #3 $\forall y \in [6;1\ \ ]$ $\quad f_{X|Y}(x,y) = \frac{1}{x \cdot \ln\frac{16}{y}}$

(d). A wins the race if and only if $x \prec y$.

If $y = \frac{3}{4}$, then $y \prec 3 \le x$ and $P\left(S \middle| y = \frac{3}{4}\right) = 0$, whereas obviously, $P(S) \neq 0$. Thus, the event S and r.v. Y are not independent.

(e). The winner will win the by less than 1 min if and only if $|x - y| \le 1$.

The corresponding area on the joint sample space is:



The probability that the winner will win by less than 1 minute is then:

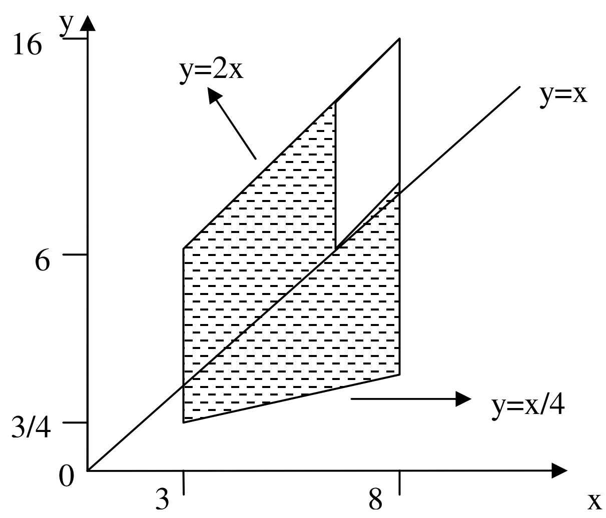
$$P = \iint_{Area} f_{X,Y}(x, y)dxdy$$

$$P = \iint_{Area} \frac{4}{35x} dxdy$$

$$P = -\frac{8\ln\left(\frac{3}{8}\right)}{35}$$

$$P \approx 0.22$$

(f). The winner's time is less than 6 minutes if and only if $\min(x, y) \le 6$. On the joint sample



space, this corresponds to:

Thus : $P = 1 - \iint_{\substack{unshaded \\ area}} f_{X,Y}(x, y)dxdy = 1 - \int_{x=6}^{8} \int_{y=x}^{y=2x} \frac{4}{35x} dxdy = 0.77$

Problem 2
Cell Phones

(a). Let i be the day of the month: $i = 1..30$.
Let N_i be the number of phone calls you make/receive during day i.
Let $T_{i,j}$ be the number of minutes you spend on the j^{th} phone call of day i, with $j = 1..N_i$.

Then, the total number of minutes you spend on the phone per month is:

$$T_{tot} = \sum_{i=1}^{30} \sum_{j=1}^{N_i} T_{i,j} \quad .$$

The N_i follow a Poisson distribution with mean 3 calls/day, and are statistically independent.
The $T_{i,j}$ follow an exponential distribution with mean 5 minutes, and are statistically independent.

Therefore,

$$E[T_{tot}] = E\left[\sum_{i=1}^{30} \sum_{j=1}^{N_i} T_{i,j}\right]$$

$$E[T_{tot}] = \sum_{i=1}^{30} E\left[\sum_{j=1}^{N_i} T_{i,j}\right]$$

$$E[T_{tot}] = \sum_{i=1}^{30} \left(E[T_{i,j}] \cdot E[N_i]\right)$$

$$E[T_{tot}] = \sum_{i=1}^{30} 5 \times 3$$

$$E[T_{tot}] = 30 \times 5 \times 3$$

$$E[T_{tot}] = 450(\min)$$

$$Var[T_{tot}] = Var\left[\sum_{i=1}^{30} \sum_{j=1}^{N_i} T_{i,j}\right]$$

$$Var[T_{tot}] = \sum_{i=1}^{30} Var\left[\sum_{j=1}^{N_i} T_{i,j}\right]$$

If $Y = X_1 \ldots X_N$, where the X_i are i.i.d. and N is a r.v., then
$Var[Y] = E[N] \cdot VarX + (E[X])^2 \cdot VarN$.
Thus:

$$Var[T_{tot}] = \sum_{i=1}^{30} \left(E[N_i] \cdot VarT_{i,j} + \left(E[T_{i.j}] \right)^2 \cdot VarN_i \right)$$
$$Var[T_{tot}] = \sum_{i=1}^{30} \left(3 \cdot 25 + 5^2 \cdot 3 \right)$$
$$Var[T_{tot}] = 30\left(3 \cdot 25 + 5^2 \cdot 3 \right)$$
$$Var[T_{tot}] = 4{,}500(\text{mi}\quad)$$

(b). Let $N_{i,j}$ be the number of phone calls you make/accept the the i^{th} of the j^{th} month of the year. We have $\begin{cases} i = 1 \ldots 30 \\ j = 1 \ldots 12 \end{cases}$.

The total number of phone calls is therefore $N_{tot} = \sum_{j=1}^{12} \sum_{i=1}^{30} N_{i,j}$.

The total number of days is 360. It is sufficiently large to apply the Central Limit Theorem. Thus N_{tot} has approximately a normal distribution with mean $\mu_{N_{tot}} = 360 \cdot E(N_{i.j}) = 360 \cdot 3 = 1080$ and variance $\sigma^2_{N_{tot}} = 360 \cdot \sigma^2_{Ni,j} = 360 \cdot 3 = 1080$.

We will now normalize the variable studied in order to read the answer from a table:

$$P(1190 \le N_{tot} \le 1210) = P\left(\frac{110}{\sqrt{1080}} \le \frac{N_{tot} - 1080}{\sqrt{1080}} \le \frac{130}{\sqrt{1080}} \right)$$
$$P(1190 \le N_{tot} \le 1210) = \Phi\left(\frac{130}{\sqrt{1080}} \right) - \Phi\left(\frac{110}{\sqrt{1080}} \right)$$
$$P(1190 \le N_{tot} \le 1210) \approx 0$$

(c). We are looking for the pdf or $T_{tot} = \sum_{i=1}^{4} T_i$, where the T_i are i.i.d. exponential with mean 5 minutes. Thus, T_{tot} has an Erlang order 4 distribution (see book p.49):

$$f_{T_4}(t) = \begin{cases} \dfrac{\lambda^4 t^3 e^{\lambda \cdot t}}{3!}, t \ge 0 \\ 0, otherwise \end{cases} \text{ and } \lambda = \frac{1}{5}\,\text{min}^{-1}.$$

(d). On another given day, you are told that you will spend exactly 20 minutes total in phone conversation time. Determine the conditional probability mass function for the number of different phone calls yielding those 20 minutes of conversation.

Let T be the number of minutes you spend in conversation time on that given day.
Let N be the number of phone calls you made or receive that day.
We are looking for $P(N = k | T = 20\,\text{min})$ for $k = 1 \ldots \infty$.

$$P(N=k|T=20\min)=\frac{P(T=20\min|N=k)\cdot P(N=k)}{P(T=20\min)}$$

$$P(N=k|T=20\min)=\frac{P(T=20\min|N=k)\cdot P(N=k)}{\sum_{i=1}^{\infty}P(T=20\min|N=i)\cdot P(N=i)}$$

We can use the same reasoning as in question c) to find $P(T=20\min|N=i)$.

Thus, for $i=1\ldots\infty$, the conditional pdf for T is $f_{T_i}(t)=\begin{cases}\dfrac{\lambda^i t^{i-1}e^{\lambda\cdot t}}{(i-1)!}, t\geq 0\\ 0, otherwise\end{cases}$.

And by definition: $P(T=20\min|N=i)=\lim_{dt\to 0}\left(f_{T_i}(t)\cdot dt\right)$. Therefore we have:

$$P(N=k|T=20\min)=\lim_{dt\to 0}\frac{f_{T_k}(20)\cdot dt\cdot P(N=k)}{\sum_{i=1}^{\infty}f_{T_k}(20)\cdot dt\cdot P(N=i)}$$

$$P(N=k|T=20\min)=\lim_{dt\to 0}\frac{f_{T_k}(20)\cdot\left(\dfrac{3^k e^{-3}}{k!}\right)}{\sum_{i=1}^{\infty}f_{T_k}(20)\cdot\left(\dfrac{3^i e^{-3}}{i!}\right)}$$

(e). From part a), we have our total time spent on the phone per month: $T_{tot}=\sum_{i=1}^{30}\sum_{j=1}^{N_i}T_{i,j}$.

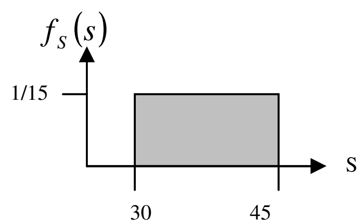
With plan i, we would pay at the end of the month: $D_i+C_i\max\left(T_{tot}-M_i,0\right)$.
We would choose the plan that minimizes the expected value of this expression.
Thus, we are interested in:
$\arg_i\min(E[D_i+C_i\max\left(T_{tot}-M_i,0\right)])=\arg_i\min\{D_i+C_i\cdot E[\max\left(T_{tot}-M_i,0\right)]\}$
We can use the Central Limit Theorem to approximate the distribution of T_{tot}. We would have to use the Central Limit Theorem again to simulate the expected value of the maximum, and then choose the appropriate plan from the analysis of the different simulations.

(f). The advocate's statement may not be true. Consider the following counterexample in which there are only 2 plans available. $(D_1,C_1,M_1)=(1,1,1)$ and $(D_2,C_2,M_2)=(99,0.5,10^5)$. Let T be the r.v. representing the customer's talk time in a month. Suppose we have that $P(T\geq M_2)=0$ and $P(T\geq 100)=1$. Then, the customer will always spend at least \$100 on plan 1 but never more than \$99 on plan 2, even though he/she never uses all M_2 minutes in plan 2. So, plan 2 is actually cheaper for this customer.

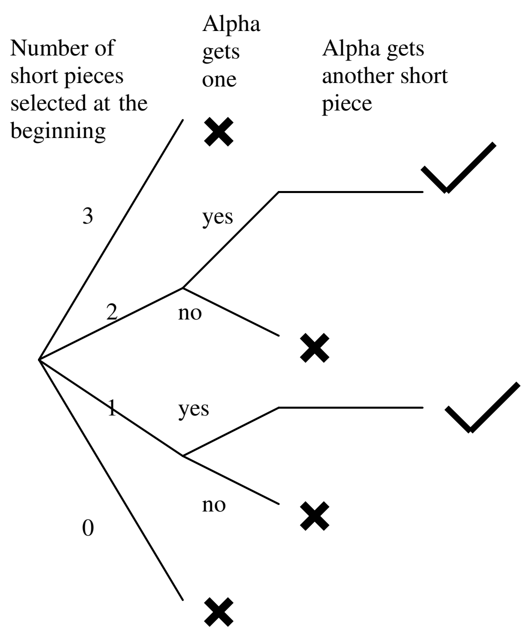
Problem 3
Dogs in the woods

a). Let S be the random variable representing the number of calories in a short piece.
We have $s \in [30;45]$ (calories).
Because the break is uniformly distributed in the [30;60] interval, the pdf of S is also uniformly distributed, but over [30;45].

Thus, the pdf we are looking for is: $f_S(s) = \frac{1}{15}$ for $s \in [30;45]$:

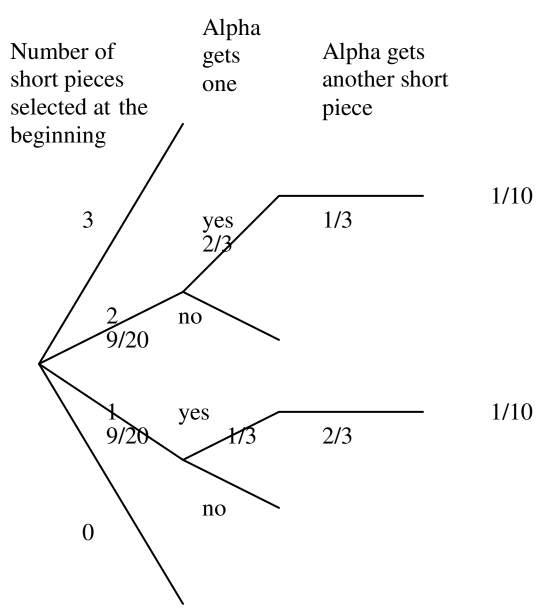


(b). Before starting running, Professor X breaks all the biscuits into two pieces. Thus, he has 3 short and 3 long pieces in his pocket at the beginning of the run.
In order to have two short pieces, Alpha should get a short piece from the six pieces at the beginning of the run, and another short piece from the three pieces remaining at the end of the run.
The decision tree for "Alpha gets exactly two short pieces" is:



	Number n of short pieces selected at the beginning	Probability that n short pieces are selected at the beginning $\frac{C_3^n \cdot C_3^{n-3}}{C_6^3}$	Probability that Alpha gets one of the short pieces	Number of short pieces remaining	Probability that Alpha will get another short piece at the end
Case #1	3	$\frac{C_3^3 \cdot C_3^0}{C_6^3} = \frac{1}{20}$	1	0	0
Case #2	2	$\frac{C_3^2 \cdot C_3^1}{C_6^3} = \frac{9}{20}$	2/3	1	$\frac{C_1^1}{C_3^1} = \frac{1}{3}$
Case #3	1	$\frac{C_3^1 \cdot C_3^2}{C_6^3} = \frac{9}{20}$	1/3	2	$\frac{C_2^1}{C_3^1} = \frac{2}{3}$
Case #4	0	$\frac{C_3^0 \cdot C_3^3}{C_6^3} = \frac{1}{20}$	0	3	1

Then, we have the information needed for the decision tree:

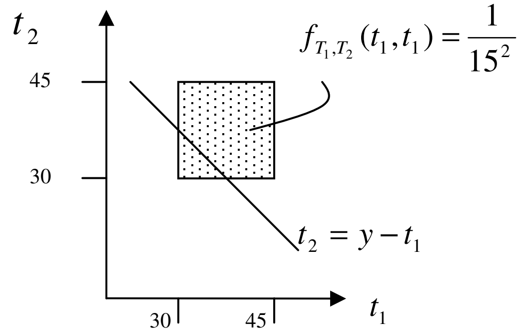


Therefore, the probability that Alpha gets two short pieces is: $P = \frac{1}{10} + \frac{1}{10} = \frac{1}{5}$.

(c). Alpha gets exactly two short pieces T_1 and T_2 that have their number of calories uniformly distributed over [30;45] and that are independent.

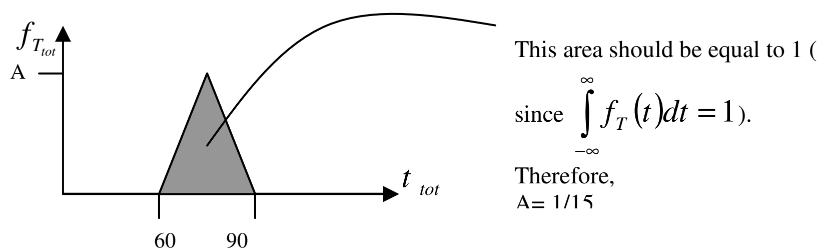
Thus: $f_{T_1,T_2}(a,b) = \dfrac{1}{15^2}$ for $(a,b) \in [30;45]^2$.

The joint sample space is therefore:



The area defined by the equation gives $f_{T_{tot}}(y)$

Therefore, the pdf of T_{tot} has the following shape:



(d). Beta will receive exactly 90 calories from biscuit pieces today if and only if he gets the two corresponding parts of a biscuit.
Since Alpha receives exactly two short pieces, there is only one short piece remaining and three long ones for Beta.

Thus the probability is : $P = \dfrac{C_2^2}{C_4^2} = \dfrac{1}{6}$.

(e). Each dog receives two pieces of biscuits and the number of calories of each piece is uniformly distributed over [30;60]. We are then considering $T_{tot} = T_1 + T_2$ where T_1 and T_2 are the number of calories from the first piece and second piece.

There are two cases: - the two pieces correspond to the same biscuit;
- the two pieces do not correspond.

The probability to pick two corresponding pieces from the 6 pieces we have is:

$$P = \frac{Number\_corresponding\_pairs}{Number\_possible\_pairs} = \frac{6}{6*5} = \frac{1}{5}.$$

If the two pieces correspond, the number of calories is exactly 90. It is deterministic.

If the two pieces do not correspond, then they are independent, and therefore, the variance is given by: $\sigma^2_{Tot,independent} = 2 \cdot \sigma^2 = 2 \cdot \frac{(60-30)^2}{12} = 150$.

The variance in daily biscuit caloric intake is therefore:

$$\sigma^2_{Tot} = \sigma^2_{Tot,independent} \cdot (1-P) + 0 * P$$

$$\sigma^2_{Tot} = \frac{4}{5} \cdot 150 = 120$$

(f). We are now considering $T_{tot} = T_{short} + T_{long}$, with T_{short} representing the number of calories in a short piece and T_{long} representing the number of calories in a long piece.

T_{short} is uniformly distributed over [30;45], and T_{long} is uniformly distributed over [45:60].

We also have two cases:

- the two parts correspond to the same biscuit, with a probability $P = \frac{6}{18} = \frac{1}{3}$.
- the two parts are from two different biscuits, thus T_{long} and T_{short} are independent.

If the two parts correspond, then the number of calories is exactly 90.

If the two parts are different, then

$$\sigma^2_{Tot,independent} = \sigma^2_{T_{long}} + \sigma^2_{T_{short}} = \frac{(60-45)^2}{12} + \frac{(45-35)^2}{12} = 37.5 .$$

Thus, for this question, the variance is:

$$\sigma^2_{Tot} = \sigma^2_{Tot,independent} \cdot (1-P) + 0 * P$$

$$\sigma^2_{Tot} = \frac{2}{3} \cdot 37.5 = 25$$

Problem 4
Pedestrian Crossing Problem, revisited

There is no single correct answer to this problem.

The goal is to design a system that takes into account both the pedestrians, who do not want to wait too long before the light turns green for them, and the drivers, who do not want to stop too often.

We are therefore mainly interested in:

1. the expected time that a randomly arriving pedestrian must wait;
2. the expected time between dumps.

One of these two measures can be fixed. Let's say we first want the expected waiting time for a pedestrian to be a constant for the three rules. The parameters of the rules (T, To and N) required to achieve such a goal could then be determined, and the other measure, the expected time between dumps, computed. That could be repeated for different values of the arrival rate of pedestrians λ. The arrival rate of cars should not be neglected either. We do not want traffic jams in front of 77 Mass. Ave.

That would be one way of comparing the different rules.

Not all the rules are easy to implement. Rule A and Rule C are more systematic and reliable than Rule B.

From that analysis, we can assign one rule to each particular situation (defined by the arrival rates of the pedestrians and the cars).

Once that we have a list of all the different types of situations, with their corresponding arrival rates and suggested rule, we can determine in which categories falls each time slot of the week. A schedule could then be proposed to our two sponsors.