Problem Set 5: Solutions
Due: March 15, 2010

1. Let random variable X be the lifetime of a randomly select chip, and define events as follows:

A_t: Chip still works at time t.
B: The chip is bad.
G: The chip is good.

Using the given exponential distributions,

$$\begin{aligned} \mathbf{P}(A_t \mid G) &= \int_t^\infty \alpha e^{-\alpha x}\,dx = e^{-\alpha t} \qquad \text{and} \\ \mathbf{P}(A_t \mid B) &= \int_t^\infty 1000\alpha e^{-1000\alpha x}\,dx = e^{-1000\alpha t}. \end{aligned}$$

(a) By using the the total probability theorem,

$$\mathbf{P}(A_t) = \mathbf{P}(G)\mathbf{P}(A_t \mid G) + \mathbf{P}(B)\mathbf{P}(A_t \mid B) = pe^{-\alpha t} + (1-p)e^{-1000\alpha t}.$$

(b) We are asked for the PDF of X, and the computation of part (a) indirectly gives the CDF of X:

$$F_X(x) = \mathbf{P}(X \le x) = 1 - \mathbf{P}(A_x) = 1 - pe^{-\alpha x} - (1-p)e^{-1000\alpha x}.$$

Thus

$$f_X(x) = \frac{d}{dx}F_X(x) = p\alpha e^{-\alpha x} + (1-p)1000\alpha e^{-1000\alpha x}.$$

(c) By using the definition of conditional probability, we get:

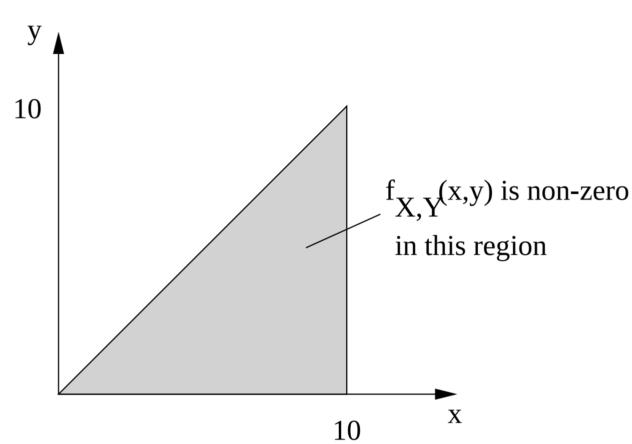
$$\mathbf{P}(B \mid A_t) = \frac{\mathbf{P}(B \cap A_t)}{\mathbf{P}(A_t)}.$$

Furthermore, $\mathbf{P}(B \cap A_t) = \mathbf{P}(B)\mathbf{P}(A_t \mid B) = (1-p)e^{-1000\alpha t}$. Therefore,

$$\mathbf{P}(B \mid A_t) = \frac{(1-p)e^{-1000\alpha t}}{pe^{-\alpha t} + (1-p)e^{-1000\alpha t}} = \frac{1}{(p/(1-p))e^{999\alpha t} + 1}.$$

To make $\mathbf{P}(B \mid A_t) < 0.01$ when $p/(1-p) = 9$, we must have $e^{999\alpha t} > 11$, or $t > (\ln 11)/(999\alpha)$.

2. Let us first find the marginal probability of X and the conditional probability of Y.



$$f_X(x) = \int_0^x f_{X,Y}(x,y)dy = \int_0^x Qx^2 y dy = Qx^2 \left[\frac{y^2}{2}\right]_0^x = \frac{Qx^4}{2},$$

Therefore

$$f_X(x) = \begin{cases} \frac{Qx^4}{2}, & 0 \le x \le 10 \\ 0, & \text{otherwise} \end{cases}$$

And for the conditional probability,

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{2Qx^2y}{Qx^4} = \frac{2y}{x^2},$$

or

$$f_{Y|X}(y|x) = \begin{cases} \frac{2y}{x^2}, & 0 \le x \le 10, 0 \le y \le x \\ 0, & \text{otherwise} \end{cases}$$

Let A be the event "Experimental value of X is 6", and B be the event "Experimental value of X is 8". Then

$$\mathbf{P}(A) = \mathbf{P}(B) = \frac{1}{2}.$$

Now to find the density of Y we can condition on whether A or B occurs. Hence, we get

$$f_y(y) = f_{Y|A}(y|A) \cdot \mathbf{P}(A) + f_{Y|B}(Y|B) \cdot \mathbf{P}(B) = f_{Y|X}(y|6) \cdot \frac{1}{2} + f_{Y|X}(y|8) \cdot \frac{1}{2}$$

From the formula derived above,

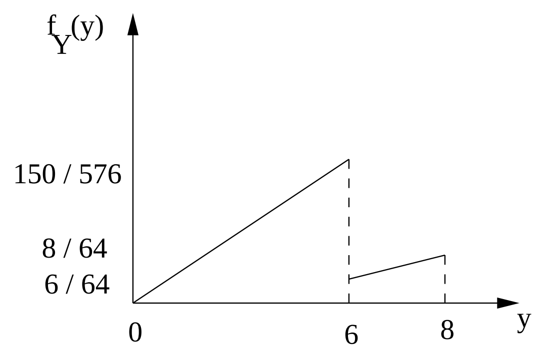
$$f_{Y|X}(y|6) = \begin{cases} \frac{y}{18}, & 0 \le y \le 6 \\ 0, & \text{otherwise} \end{cases}$$

$$f_{Y|X}(y|8) = \begin{cases} \frac{y}{32}, & 0 \le y \le 8 \\ 0, & \text{otherwise} \end{cases}$$

This implies

$$f_Y(y) = \begin{cases} \frac{25y}{576}, & 0 \le y \le 6 \\ \frac{y}{64}, & 6 < y \le 8 \\ 0, & \text{otherwise} \end{cases}$$

Graphical representation:



3. We compute the PMF of N by finding the joint PMF of N and p and then proceeding to find the marginal density–in this case the PMF:

$$
\begin{aligned}
\mathbf{P}(N=k) &= \int_p \mathbf{P}(N=k, \mathbf{P}=p)\, dp \\
&= \int_p \mathbf{P}(N=k \mid \mathbf{P}=p) \cdot f_P(p)\, dp \\
&= \int_0^{1-1/n} (1-p)^{k-1} p \frac{n}{n-1}\, dp \text{ let } q = 1-p \\
&= \frac{n}{n-1} \int_1^{1/n} (q^{k-1} - q^k) \; - dq \\
&= \frac{n}{n-1} \int_{1/n}^{1} (q^{k-1} - q^k)\, dq \\
&= \frac{n}{n-1} \left[\frac{q^k}{k} - \frac{q^{k+1}}{k+1} \right]_{1/n}^{1} \\
&= \frac{n}{n-1} \left[\frac{1 - (\frac{1}{n})^k (k+1) - (\frac{1}{n})^{k+1} k}{k^2+k} \right]
\end{aligned}
$$

and finally we have:

$$\lim_{n\to\infty} \mathbf{P}(N=k) = \frac{1}{k^2+k}$$

which is easily checked by finding $\mathbf{P}(N=k)$ where N is a geometric random variable with parameter p, and p is uniformly distributed from 0 to 1.

4. (a) We know that the total length of the edge for red interval is two times that for black inteval. Since the ball is equally likely to fall in any position of the edge, probability of falling in a red interval is $\mathbf{P}(R) = \frac{2}{3}$.

   (b) Conditioned on the ball having fallen in a black interval, the ball is equally likely to fall anywhere in the interval. Thus, the PDF is

$$f_{Z|B}(z) = \begin{cases} \frac{15}{\pi r} & , \quad z \in \left[0, \frac{\pi r}{15}\right] \\ 0 & , \quad \text{otherwise} \end{cases}$$

   (c) We can find the conditional PDF of Z when the ball lands on a red interval as we did in the previous part when the ball lands on a black interval.

$$f_{Z|R}(z) = \begin{cases} \frac{15}{2\pi r} & , \quad z \in \left[0, \frac{2\pi r}{15}\right] \\ 0 & , \quad \text{otherwise} \end{cases}$$

   Using the total probability theorem, $f_Z(z) = f_{Z|R}(z)\mathbf{P}(R) + f_{Z|B}(z)\mathbf{P}(B)$. Combining this with the results of the previous parts and noting that $\mathbf{P}(B) = \frac{1}{3}$,

$$f_Z(z) = \begin{cases} \frac{10}{\pi r} & , \quad z \in \left[0, \frac{\pi r}{15}\right) \\ \frac{5}{\pi r} & , \quad z \in \left[\frac{\pi r}{15}, \frac{2\pi r}{15}\right] \\ 0 & , \quad \text{otherwise} \end{cases}$$

(d) The total gains (or losses), T, equals to the sum of all X_i, i.e. $T = X_1 + X_2 + \cdots + X_n$, Since all the X_i's are independent of each other, and they have the same Gaussian distribution, the sum will also be a Gaussian with

$$\mathbf{E}[T] = \mathbf{E}[X_1] + \mathbf{E}[X_2] + \cdots + \mathbf{E}[X_n] = 0$$

$$\text{var}(T) = \text{var}(X_1) + \text{var}(X_2) + \cdots + \text{var}(X_n) = n\sigma^2$$

Therefore, the standard deviation for T is $\sqrt{n}\sigma$.

(e)

$$\begin{aligned}
\mathbf{P}(|T| > 2\sqrt{n}\sigma) &= \mathbf{P}(T > 2\sqrt{n}\sigma) + \mathbf{P}(T < -2\sqrt{n}\sigma) \\
&= 2\mathbf{P}(T > 2\sqrt{n}\sigma) \\
&= 2\left(1 - \Phi\left(\frac{2\sqrt{n}\sigma - \mathbf{E}[T]}{\sigma_T}\right)\right) \\
&= 2(1 - \Phi(2)) \simeq 0.0454.
\end{aligned}$$

5. (a) Let A be the event that the machine is functional. Conditioned on the random variable Q taking on a particular value q, $\mathbf{P}(A|Q = q) = q$. Using the continuous form of the total probability theorem, the probability of event A is given by:

$$\begin{aligned}
\mathbf{P}(A) &= \int_0^1 \mathbf{P}(A|Q = q) f_Q(q) dq \\
&= \int_0^1 q\, dq \\
&= 1/2
\end{aligned}$$

(b) Let B be the event that the machine is functional on m out of the last n days. Conditioned on random variable Q taking on value q (a probability q of being functional) the probability of event B is binomial with n trials, m successes, and a probability q of success in each trial. Again using the total probability theorem, the probability of event B is given by:

$$\begin{aligned}
\mathbf{P}(B) &= \int_0^1 \mathbf{P}(B|Q = q) f_Q(q) dq \\
&= \int_0^1 \binom{n}{m} q^m (1-q)^{n-m} f_Q(q) dq \\
&= \binom{n}{m} \frac{m!(n-m)!}{(n+1)!}
\end{aligned}$$

We then find the distribution on Q conditioned on event B using Bayes rule:

$$\begin{aligned}
f_{Q|B}(q) &= \frac{\mathbf{P}(B|Q = q) f_Q(q)}{\mathbf{P}(B)} \\
&= \frac{q^m(1-q)^{n-m}}{\frac{m!(n-m)!}{(n+1)!}} \quad 0 \le q \le 1, \quad n \ge m
\end{aligned}$$

(c) Let event C be the probability that the machine is functional today. The probability $\mathbf{P}(C|B)$ is given by:

$$\begin{aligned}
\mathbf{P}(C|B) &= \int_0^1 \mathbf{P}(C|Q=q,B) f_{Q|B}(q)dq \\
&= \int_0^1 \mathbf{P}(C|Q=q) f_{Q|B}(q)dq \\
&= \int_0^1 \frac{q \cdot q^m (1-q)^{n-m}}{\frac{m!(n-m)!}{(n+1)!}} dq \quad n \geq m \\
&= \int_0^1 \frac{q^{m+1} (1-q)^{(n+1)-(m+1)}}{\frac{m!(n-m)!}{(n+1)!}} dq \quad n \geq m \\
&= \frac{\frac{(m+1)!(n-m)!}{(n+2)!}}{\frac{m!(n-m)!}{(n+1)!}} \quad n \geq m \\
&= \frac{m+1}{n+2} \quad n \geq m
\end{aligned}$$

where the second equality follows since events C and B are conditionally independent given Q.

G1†. Let X_1, X_2, ..., X_n be independent, identically distributed (IID) random variables.

We note that

$$\mathbf{E}[X_1 + \cdots + X_n | X_1 + \cdots + X_n = x] = x.$$

It follows from the linearity of expectations that

$$\begin{aligned}
x &= \mathbf{E}[X_1 + \cdots + X_n | X_1 + \cdots + X_n = x] \\
&= \mathbf{E}[X_1 | X_1 + \cdots + X_n = x] + \cdots + \mathbf{E}[X_n | X_1 + \cdots + X_n = x]
\end{aligned}$$

Because the X_i's are identically distributed, we have the following relationship.

$$\mathbf{E}[X_i | X_1 + \cdots + X_n = x] = \mathbf{E}[X_j | X_1 + \cdots + X_n = x], \text{ for any } 1 \leq i \leq n, 1 \leq j \leq n.$$

Therefore,

$$\begin{aligned}
n\mathbf{E}[X_1 | X_1 + \cdots + X_n = x] &= x \\
\mathbf{E}[X_1 | X_1 + \cdots + X_n = x] &= \frac{x}{n}.
\end{aligned}$$

†Required for 6.431; optional for 6.041