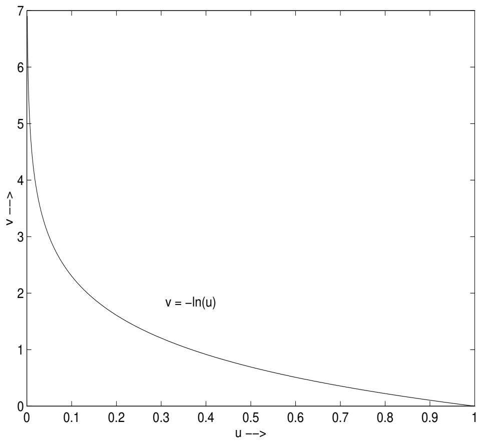
University
of Illinois

Solutions: Problem Set 10

ECE 313
Fall 1998

1. We have a random variable $\mathbf{X} \sim U(0,1)$.

   (a) Define the RV $\mathbf{Y} = -\ln(\mathbf{X})$. This is clearly a continuous RV, and hence has a pdf. The graph of $v = -\ln(u)$ is shown below.



As can be seen from the graph of $-\ln(u)$, the random variable $\mathbf{Y}$ takes on all values in $(0, \infty)$.

The pdf of $\mathbf{Y}$ can be given as $f_Y(v) = f_X(u_1)/|g'(u_1)|$, where u_1 is the single root of the equation $v = -\ln(u)$, and is equal to $\exp(-v)$. Accordingly, we have $g'(u_1) = -\exp(v)$ and

$$f_Y(v) = \begin{cases} e^{-v}, & v \geq 0 \\ 0\,, & v < 0 \end{cases}$$

   (b) You don't need to calculate the pdf of $\sqrt{\mathbf{X}}$ for this part! Let $\mathbf{Z}_2$ be the RV corresponding to the second decimal digit of the square root of $\mathbf{X}$, and $\mathbf{Z}_1$ be the RV corresponding to its first digit. Both $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are discrete RVs, taking on values in $\{0, 1, \ldots, 9\}$. If $\mathbf{Z}_1 = m$, then the probability that $\mathbf{Z}_2 = k$ is given by the probability that $\{10m + k \leq 100\sqrt{\mathbf{X}} < 10m + k + 1\}$ (why is this true?). Therefore, the probability that $\mathbf{Z}_1 = m$ and $\mathbf{Z}_2 = k$ is obtained as

$$\begin{aligned} P\{\mathbf{Z}_1 = m, \mathbf{Z}_2 = k\} &= P\{10m + k \leq 100\sqrt{\mathbf{X}} < 10m + k + 1\} \\ &= P\left\{\frac{1}{10^4}(10m+k)^2 \leq \mathbf{X} < \frac{1}{10^4}(10m+k+1)^2\right\} \\ &\overset{a}{=} \frac{1}{10^4}\big((10m+k+1)^2 - (10m+k)^2\big) = \frac{20m + 2k + 1}{10000} \end{aligned}$$

Summing over all values that $\mathbf{Z}_1$ can take, we get the pmf of $\mathbf{Z}_2$ as

$$p_{Z_2}(k) = \sum_{m=0}^{9} \frac{20m + 2k + 1}{10000} = 0.091 + 0.002k.$$

(Check and see that this sums to 1 over all values of k.)

2. (a) If $\mathbf{Y} = \mathrm{sign}(\mathbf{X}+2)$, then $\mathbf{Y}$ is a discrete random variable taking on 2 values — $+1$ and -1. $\mathbf{Y} = +1$ if $\mathbf{X} \geq -2$ and $\mathbf{Y} = -1$ if $\mathbf{X} < -2$. Therefore

$$p_Y(-1) = P\{\mathbf{X} \geq -2\} = \int_{-\infty}^{-2} \frac{1}{2w^2}\,dw = \left(\frac{-w^{-1}}{2}\right)\Bigg|_{-\infty}^{-2} = \frac{1}{4}; \qquad p_Y(+1) = 1 - p_Y(-1) = \frac{3}{4}$$

(b) Let us find $F_Y(v)$, the CDF of $\mathbf{Y}$. If v represents the value for $\mathbf{Y}$ and u represents the value for $\mathbf{X}$, then the figure below (Fig 1) graphs $v = g(u)$, where $g(u) = 2|u|$, for $|u| \leq 2$; and $v = u^2$, for $|u| \geq 2$. First of all, we see that v takes only positive values. Furthermore, $v \geq 2$, because $P\{-1 \leq \mathbf{X} \leq 1\} = 0$.

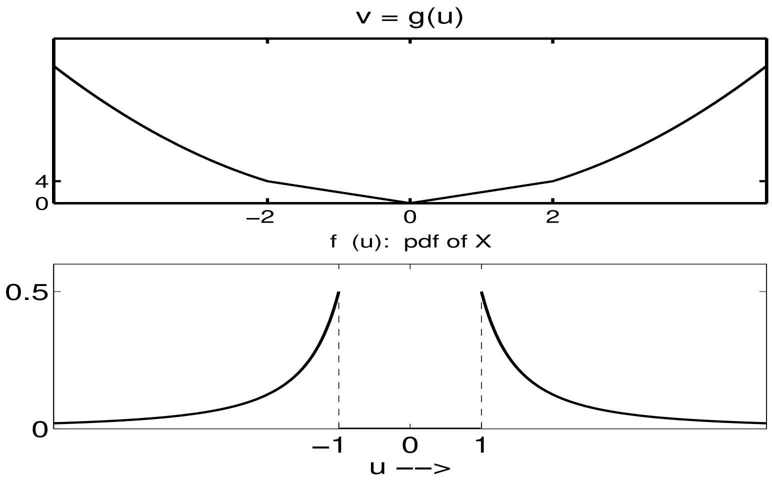


Figure 1: Top plot: $v = g(u)$, where v and u represent typical values of random variables $\mathbf{Y}$ and $\mathbf{X}$ respectively. Bottom plot: pdf of $\mathbf{X}$.

i. If $v \leq 2$: $P\{\mathbf{Y} \leq v\} = 0$, and so $F_Y(v) = 0$.

ii. If $2 \leq v \leq 4$:

$$P\{\mathbf{Y} \leq v\} = P\left\{-\frac{v}{2} \leq \mathbf{X} \leq \frac{v}{2}\right\} = P\left\{(-\frac{v}{2} \leq \mathbf{X} \leq -1) \cup (1 \leq \mathbf{X} \leq \frac{v}{2})\right\}$$
$$= F_X(v/2) - F_X(1) + F_X(-1) - F_X(-v/2) = F_X(v/2) - F_X(-v/2)$$

since $F_X(1) = F_X(-1)\ \ (= 1/2)$.

iii. If $v > 4$:

$$P\{\mathbf{Y} \leq v\} = P\{-\sqrt{v} \leq \mathbf{X} \leq \sqrt{v}\} = F_X(\sqrt{v}) - F_X(-\sqrt{v})$$

by reasoning similar to the above.

Differentiating $F_Y(v)$ gives us the pdf of $\mathbf{Y}$. Therefore,

$$f_Y(v) = \begin{cases} 0, & v < 2 \\ \frac{1}{2}\big(f_X(v/2) + f_X(-v/2)\big) = \frac{2}{v^2}, & 2 \leq v < 4 \\ \frac{1}{2\sqrt{v}}\left(f_X(\sqrt{v}) + f_X(-\sqrt{v})\right) = \frac{1}{2v^{3/2}}, & v \geq 4 \end{cases}$$

3. $\mathbf{X}$ is a continuous RV with CDF $F_X(u)$.

   (a) Because $\mathbf{Y} = F_X(\mathbf{X})$, we know it is bounded between 0 and 1.

   $$F_Y(v) = P\{\mathbf{Y} \leq v\} = P\{F_X(\mathbf{X}) \leq v\} \stackrel{(a)}{=} P\{\mathbf{X} \leq F_X^{-1}(v)\} = F_X\left(F_X^{-1}(v)\right) \stackrel{(b)}{=} v.$$

   Therefore $\mathbf{Y}$ is a uniform RV in the interval $(0, 1)$. Equalities (a) and (b) follow because $F_X(u)$ is a monotonic nondecreasing function and hence, one-to-one. Hence, for all values of $v \in (0, 1)$, $F_X^{-1}(v)$ is well defined; also, $F_X(F_X^{-1}(v)) = v$ because of the existence of the inverse.

   (b) $F_X(u) = P\{\mathbf{X} \leq u\} = P\{\mathcal{F}^{-1}(\mathbf{Y}) \leq u\} \stackrel{(a)}{=} P\{\mathbf{Y} \leq \mathcal{F}(u)\} \stackrel{(b)}{=} \mathcal{F}(u)$. Equality (a) follows because $\mathcal{F}$ is monotone nondecreasing — this is why the inequality "$\leq$" is preserved. Equality (b) follows because $\mathbf{Y}$ is a uniform RV in $[0, 1]$; care must be taken, though, to ensure that $\mathcal{F}(u) \in [0, 1]$, but since $\mathcal{F}$ is a CDF, this holds trivially for all u. Therefore we have generated a RV $\mathbf{X}$ with the desired CDF $\mathcal{F}(u)$ for all u.

4. The new RV $\mathbf{R}$, which is the received signal, is defined by $\mathbf{R} = s + \mathbf{X}$.

   (a) $\mathbf{R}$ is a continuous RV with pdf $f_R(r) = \dfrac{1}{\sigma\sqrt{2\pi}} \exp\left(-\dfrac{(r-s)^2}{2\sigma^2}\right)$ for all $-\infty < r < \infty$. Thus in fact, $\mathbf{R}$ is a Gaussian RV with mean s and variance σ^2.

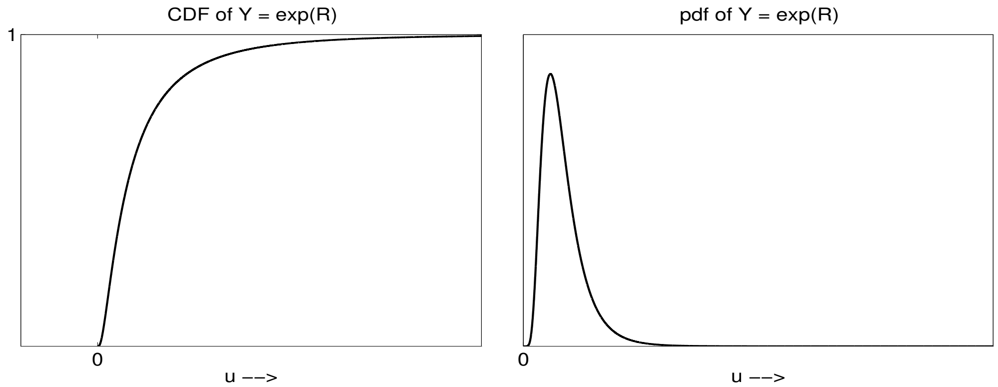
   (b) Define a new RV $\mathbf{Y} = e^{\mathbf{R}}$. The function $v = g(r) \triangleq e^r$ has exactly one root, $r_1 = \ln v$, for $v > 0$ and no roots for $v \leq 0$. Thus,

   $$\begin{aligned} f_Y(v) &= \frac{f_R(r_1)}{|g'(r_1)|} = \frac{1}{\sigma\sqrt{2\pi}\cdot v} \exp\left(-\frac{(\ln v - s)^2}{2\sigma^2}\right), \qquad v > 0 \\ &= 0, \qquad v \leq 0 \end{aligned}$$

   and the CDF is given by

   $$F_Y(v) = P(\mathbf{Y} \leq v) = P(e^{\mathbf{R}} \leq v) = P(\mathbf{R} \leq \ln v) = \begin{cases} \Phi\left(\dfrac{\ln v - s}{\sigma}\right), & v > 0 \\ 0, & v \leq 0 \end{cases},$$

   where the last equality follows from the fact that $\mathbf{R} \sim N(s, \sigma^2)$. The lognormal RV, whose typical pdf and CDF are plotted below, is used to model the amplitude of the received Gaussian signal in the presence of, what is termed as, shadowing.



(c) Quantizers: The trick here is to realize that since the quantizers quantize the received signal about its mean, you can treat the received signal as a **zero mean** signal — you don't have to worry about its mean, and this makes for easier book-keeping while writing out integrals. Therefore, the quantizers can be considered to be centered about 0 with $\mathbf{X} \sim N(0, \sigma^2)$ as the input.

i. $\mathbf{Y}$ is a discrete RV in both cases, with 3 and 4 values respectively for Quantizers A and B. The pmf of $\mathbf{Y}$ for both quantizers is given below

Quantizer A	Quantizer B
$p_Y(1.5) = P\{\mathbf{X} > +c\} = 1 - \Phi(c/\sigma)$	$p_Y(1.5) = P\{\mathbf{X} > +c\} = 1 - \Phi(c/\sigma)$
$p_Y(0) = P\{-c < \mathbf{X} \le +c\} = 2\Phi(c/\sigma) - 1$	$p_Y(0.5) = P\{0 < \mathbf{X} \le +c\} = \Phi(c/\sigma) - 1/2$
	$p_Y(-0.5) = P\{-c < \mathbf{X} \le 0\} = 1/2 - \Phi(c/\sigma)$
$p_Y(-1.5) = P\{\mathbf{X} \le -c\} = 1 - \Phi(c/\sigma)$	$p_Y(-1.5) = P\{\mathbf{X} \le -c\} = 1 - \Phi(c/\sigma)$

ii. To get the value for c we need to carry out the following minimization: $\min_c E[(\mathbf{R} - \mathbf{Y})^2]$, where the expectation is with respect to the pdf of $\mathbf{R}$ (this is because $\mathbf{Y}$ is a function of $\mathbf{R}$, and hence, $(\mathbf{R} - \mathbf{Y})^2$ is a function of $\mathbf{R}$ — now we use LOTUS). since the quantizer is centered about $E[\mathbf{R}]$, we can consider a zero-mean version of $\mathbf{R}$ — call it $\mathbf{X}$. Now c^* is calculated as the value of c that minimizes $E[(\mathbf{X} - \mathbf{Y})^2]$. For Quantizer A,

$$E[(\mathbf{X} - \mathbf{Y})^2] = \frac{1}{\sigma\sqrt{2\pi}} \left(\int_{-\infty}^{-c} (u + 1.5)^2 e^{-\frac{u^2}{2\sigma^2}}\, du + \int_{-c}^{+c} u^2 e^{-\frac{u^2}{2\sigma^2}}\, du + \int_{+c}^{\infty} (u - 1.5)^2 e^{-\frac{u^2}{2\sigma^2}}\, du \right)$$

Differentiating with respect to c using the fundamental theorem of calculus, and equating to 0, we get

$$-(-c + 1.5)^2 e^{-\frac{(-c)^2}{2\sigma^2}} + c^2 e^{-\frac{c^2}{2\sigma^2}} - (-1)\cdot(-c)^2 e^{-\frac{(-c)^2}{2\sigma^2}} - (c - 1.5)^2 e^{-\frac{c^2}{2\sigma^2}} = 0$$

Since $\exp(-\frac{c^2}{2\sigma^2}) \neq 0$, we get $-2(c - 1.5)^2 + 2c^2 = 0$, from whence we get $c^* = 1.5/2 = 0.75$. For Quantizer B we can carry out similar a similar integration and obtain $c^* = (1.5 + 0.5)/2 = 1$. (We will have 4 separate areas of integration instead of the 3 for Quantizer A.)

The mean squared errors for the two quantizers (which weren't asked for) can also be calculated, and you will observe that the mean squared error for Quantizer B is less than that for Quantizer A.

5. (a) $\mathbf{X}$ takes on values from $\{2, 3, \dots, N+1\}$. $\mathbf{X}$ will take on the value k if the previous $k-1$ colors are all distinct and the k^{th} color is the same as one of the previous $k-1$. Each pick is an independent trial with the probability of any color being picked equal to $\frac{1}{N}$ (this is because there are an *equal number* of m&m's of each color — the number doesn't matter — **and** it is sampling *with* replacement). The number of ways of *arranging* $k-1$ different colors out of N is $\frac{N!}{(N-(k-1))!}$. Therefore, the pmf of $\mathbf{X}$ is given by

$$p_X(k) = \begin{cases} \dfrac{N!\cdot(k-1)}{(N-k+1)!\cdot N^k}, & k \in \{2, 3, \dots, N+1\} \\ 0, & \text{elsewhere} \end{cases}$$

(b) The minimum value that the RV $\mathbf{Y}$ takes is N, and it takes on values from the set $N, N+1, N+2, \dots$ to ∞. Let us map the N colors to the numbers $1, 2, \dots, N$. If $\mathbf{Y} = k$, $(k \geq N)$ this means that one color, say N, occurred for the first time in the k^{th} pick, and all the other $N-1$ colors each occurred at least once in the previous $k-1$ picks. Let A_i denote the event that color-i does **not** occur in the first $(k-1)$ picks. Another way of looking at the desired event now is

(i) One of the N colors (say N) occurred in the k^{th} pick — call this event $\mathcal{B}$, AND

(ii) None of the events $A_1, A_2, \dots, A_{N-1}$ occurred in the first $k-1$ trials, i.e., $\mathcal{A} = (A_1 \cup A_2 \cup \cdots \cup A_{N-1})^c$.
(iii) Get rid of the conditioning on color-N (TTP).
The probability of the event described in (i) and (ii) above is now $P(\mathcal{AB})$. This can be written as

$$P(\mathcal{AB}) = P(\mathcal{B}) - P(\mathcal{A}^c|\mathcal{B})\, P(\mathcal{B}) = P(\mathcal{B}) - P(A_1 \cup A_2 \cup \cdots \cup A_{N-1}|\mathcal{B})\, P(\mathcal{B})$$

$P(\mathcal{B}) = \frac{(N-1)^{k-1}}{N^k}$ (this is just the number of ways that $N-1$ colors can be arranged in $k-1$ slots divided by the size of the sample space, N^k), and $P(A_1 \cup \cdots \cup A_{N-1}|\mathcal{B})$ can be calculated using Prop 4.4 from Chapter 2 of Ross.

$$\begin{aligned} P(A_1 \cup \cdots \cup A_{N-1}|\mathcal{B}) &= \sum_{i=1}^{N-1} P(A_i) - \sum_{i_1, i_2} P(A_{i_1} A_{i_2}) + \ldots + (-1)^{N-2} \sum_{i_1, \dots, i_{N-2}} P(A_{i_1} A_{i_2} \cdots A_{i_{N-2}}) \\ &= \frac{\binom{N-1}{1}(N-2)^{k-1} - \binom{N-1}{2}(N-3)^{k-1} + \binom{N-1}{3}(N-4)^{k-1} - \cdots + (-1)^{N-2} 1^{k-1}}{N^k} \end{aligned}$$

Therefore

$$P(\mathcal{AB}) = \frac{(N-1)^{k-1}}{N^k} - \sum_{i=2}^{N-1} (-1)^{i-1} \binom{N-1}{i-1} \frac{(N-i)^{k-1}}{N^k} = \sum_{i=1}^{N-1} (-1)^{i-1} \binom{N-1}{i-1} \frac{(N-i)^{k-1}}{N^k}$$

The probability of the desired event ($= p_Y(k)$) is now obtained by removing the condition on $\mathcal{B}$ (i.e., that the k^{th} color is color-N). This is done by just multiplying the above probability by N, the number of possible colors that can go into the k^{th} slot. So now, $P(\mathcal{A}) = \displaystyle\sum_{i=1}^{N-1} (-1)^{i-1} \binom{N-1}{i-1} \frac{(N-i)^{k-1}}{N^{k-1}}$.
We can calculate the above for $N = 2$ and $N = 3$ to get an idea of what is going on:

<u>$N = 2$:</u> The probability that event A_2 occurs for the first time in the k^{th} trial is $\dfrac{1}{2^{k-1}} \cdot \dfrac{1}{2} = \dfrac{1}{2^k}$, and since the k^{th} trial could just as well have been A_1 (i.e., two possibilities), $P(\mathbf{Y} = k) = 2 \times \dfrac{1}{2^k} = \dfrac{1}{2^{k-1}}$.
<u>$N = 3$:</u> If A_3 is the event occurring for the first time in the k^{th} trial, the previous $k-1$ trials should only have A_1 and A_2 in them. The total number of possibilities is $2^{k-1} - 2$, where we subtract 2 for the two outcomes $A_1, A_1, \dots, A_1$ and $A_2, A_2, \dots, A_2$, both of which are clearly not allowed. Again, the k^{th} trial could be any of the 3 events. Thus, $P(\mathbf{Y} = k) = 3 \times \dfrac{2^{k-1} - 2}{3^k} = \dfrac{2^{k-1} - 2}{3^{k-1}}$.

(c) **Extra Credit:** I do not know the solution to the extra credit to part (a) (if someone has a good solution, I'd be very interested in seeing it!).

For part (b), it is best to get $E[\mathbf{Y}]$ without making use of $p_Y(k)$. This might seem odd, so lets how this works. Let X_k be the average number of picks needed to see k colors. Then X_{k+1}, the average number of picks needed to see $k+1$ colors, takes on values $X_k+1, X_k+2, \ldots, X_k+M, \ldots$. The probability that $X_{k+1} = X_k + 1$ is just the probability that one of the unseen $N-k$ colors occurs on the next pick, and this is $\frac{N-k}{N}$. Similarly, $P(X_{k+1} = X_k + 2) = \frac{K}{N} \cdot \frac{N-k}{N}$, and so on. Therefore,

$$X_{k+1} = (X_k+1)\,\frac{N-k}{N} \;+\; (X_k+2)\left(\frac{k}{N}\right)\frac{N-k}{N} \;+\; (X_k+3)\left(\frac{k}{N}\right)^2\frac{N-k}{N} + \cdots$$

$$= X_k \cdot \left(\frac{N-k}{N}\right)\left[1 + \frac{k}{N} + \left(\frac{k}{N}\right)^2 + \cdots\right] + \frac{N-k}{N}\left[1 + 2\left(\frac{k}{N}\right) + 3\left(\frac{k}{N}\right)^2 + \cdots\right]$$

The first series above is just a geometric series, summing to $\frac{1}{1-k/N} = \frac{N}{N-k}$. The second series is just the derivative of the first geometric series, and is therefore just equal to $\frac{1}{(1-k/N)^2} = \frac{N^2}{(N-K)^2}$. Plugging these values into the above equation, we get

$$X_{k+1} = X_k\left(\frac{N-k}{N}\right)\cdot\frac{N}{N-k} + \left(\frac{N-k}{N}\right)\cdot\frac{N^2}{(N-k)^2} \;=\; X_k + \frac{N}{N-k}$$

This is a recursive equation for the **average** number of picks needed to see k colors. The initial condition is $X_1 = 1$, clearly. We can construct X_N (which is what we want) as

$$X_N = N\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{N}\right).$$