Massachusetts Institute of Technology

Department of Electrical Engineering & Computer Science

6.041/6.431: Probabilistic Systems Analysis

(Fall 2011)

Problem Set 10 Solutions

1. (a) Let C denote the coin that Bob received, so that $C = 1$ if Bob received the first coin, and $C = 2$ if Bob received the second coin. Then $\mathbf{P}(C = 1) = p$ and $\mathbf{P}(C = 2) = 1 - p$. Given C, the number of heads Y in 3 independent tosses is a binomial random variable.

   We can find the probability that Bob received the first coin given that he observed k heads using Bayes' rule.

$$
\begin{aligned}
\mathbf{P}(C = 1 \mid Y = k) &= \frac{\mathbf{P}(Y = k \mid C = 1) \cdot \mathbf{P}(C = 1)}{P(Y = k \mid C = 1) \cdot \mathbf{P}(C = 1) + \mathbf{P}(Y = k \mid C = 2) \cdot \mathbf{P}(C = 2)} \\
&= \frac{\binom{3}{k} \cdot (1/3)^k (2/3)^{3-k} p}{\binom{3}{k} \cdot (1/3)^k (2/3)^{3-k} \cdot p + \binom{3}{k} \cdot (2/3)^k (1/3)^{3-k} \cdot (1-p)} \\
&= \frac{2^{3-k} p}{2^{3-k} p + 2^k (1-p)} = \frac{1}{1 + \frac{1-p}{p} 2^{2k-3}}
\end{aligned}
$$

   (b) We want to find k so that the following inequality holds.

$$
\begin{aligned}
\mathbf{P}(C = 1 \mid Y = k) &> p \\
\frac{2^{3-k} p}{2^{3-k} p + 2^k (1-p)} &> p
\end{aligned}
$$

   Note that if $p = 0$ or $p = 1$, there is no value of k that satisfies the inequality. We now solve it for $0 < p < 1$:

$$
\begin{aligned}
\frac{2^{3-k}}{2^{3-k} p + 2^k (1-p)} &> 1 \\
2^{3-k} &> 2^{3-k} p + 2^k (1-p) \\
2^{3-k}(1-p) &> 2^k (1-p) \\
2^{3-k} &> 2^k \\
2k &< 3 \\
k &< 3/2
\end{aligned}
$$

   For $0 < p < 1$, $k = 0$ or $k = 1$ the probability that Alice sent the first coin increases. The inequality does not depend on p, and so does not change when p increases. Intuitively, this makes sense: lower values of k increase Bob's belief he got the coin with lower probability of heads.

   (c) Given that Bob observes k heads, Bob must decide on whether the first or second coin was used. To minimize the error, he should decide it is the first coin when $\mathbf{P}(C = 1 \mid Y = k) \geq \mathbf{P}(C = 2 \mid Y = k)$. Thus, we have the decision rule given by

$$
\begin{aligned}
\mathbf{P}(C=1 \mid Y=k) &\geq \mathbf{P}(C=2 \mid Y=k) \\
\frac{2^{3-k}p}{2^{3-k}p+2^k(1-p)} &\geq \frac{2^k(1-p)}{2^{3-k}p+2^k(1-p)} \\
2^{3-k}p &\geq 2^k(1-p) \\
2^{2k-3} &\leq \frac{p}{1-p} \\
k &\leq \frac{3}{2}+\frac{1}{2}\log_2\frac{p}{1-p}
\end{aligned}
$$

(d) i. If $p = 2/3$, the threshold in the rule above is equal to $\frac{3+\log_2 2}{2} = 2$. Therefore, Bob will decide that he received the first coin when he observes 0, 1 or 2 heads, and will decide that he received the second coin when he observes 3 heads.
We find the probability of a correct decision using the total probability law:

$$
\begin{aligned}
\mathbf{P}(\text{Correct}) &= \mathbf{P}(\text{Correct} \mid C=1)\cdot p + \mathbf{P}(\text{Correct} \mid C=2)\cdot(1-p) \\
&= \mathbf{P}(Y<3 \mid C=1)\cdot p + \mathbf{P}(Y=3 \mid C=2)\cdot(1-p) \\
&= (1-\mathbf{P}(Y=3 \mid C=1))\cdot p + \mathbf{P}(Y=3 \mid C=2)\cdot(1-p) \\
&= (1-(1/3)^3)(2/3) + (2/3)^3(1/3) = 20/27 \approx .741
\end{aligned}
$$

ii. In absence of any data, all Bob can do is decide he received the first coin with some probability q. Note that this rule includes the deterministic decisions that he received either the first coin ($q = 1$) or the second coin ($q = 0$).
In this case, the probability of correct decision is equal to

$$
\begin{aligned}
\mathbf{P}(\text{Correct}) &= \mathbf{P}(\text{Correct} \mid C=1)\cdot p + \mathbf{P}(\text{Correct} \mid C=2)\cdot(1-p) \\
&= qp + (1-q)(1-p) = 1-p+q(2p-1) = \frac{1+q}{3}
\end{aligned}
$$

Clearly, the probability of the correct decision is maximized (or the probability of error is minimized) when $q = 1$, i.e., when Bob deterministically decides he received the first coin. In this case, $\mathbf{P}(\text{Correct}) = 2/3 \approx .667$. Observing 3 coin tosses increases the probability of the correct decision by $2/27 \approx .074$.

(e) If p is increased, the threshold in the decision rule in part (c) goes up, i.e., the range of values of k for which Bob decides he received the first coin can only go up.

(f) Bob will never decide he received the first coin if the threshold in the rule above is below zero:

$$
\begin{aligned}
\frac{3}{2}+\frac{1}{2}\log_2\frac{p}{1-p} &< 0 \\
\log_2\frac{p}{1-p} &< -3 \\
\frac{p}{1-p} &< \frac{1}{8} \\
p &< \frac{1}{9}
\end{aligned}
$$

If $p < 1/9$, the prior probability of receiving the first coin is so low that no amount of evidence from 3 tosses of the coin will make Bob decide he received the first coin.

(g) Bob will always decide he received the first coin if the threshold in the rule above is equal to or above 3:

$$\begin{aligned} \frac{3}{2} + \frac{1}{2}\log_2 \frac{p}{1-p} &\geq 3 \\ \log_2 \frac{p}{1-p} &\geq 3 \\ \frac{p}{1-p} &\geq 8 \\ p &\geq \frac{8}{9} \end{aligned}$$

If $p \geq 8/9$, the prior probability of receiving the first coin is so high that no amount of evidence from 3 tosses of the coin will make Bob decide he received the second coin.

2. (a) To find the normalization constant c we integrate the joint PDF:

$$\int_0^1 \int_0^1 f_{X,Y}(x,y)\,dy\,dx = c\int_0^1 \int_0^1 xy\,dy\,dx = c\int_0^1 1/2x\,dx = c/4.$$

Therefore, $c = 4$.

(b) To construct the conditional expectation estimator, we need to find the conditional probability density.

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{4xy}{\int_0^1 4xy\,dx} = \frac{4xy}{2y} = 2x, \quad x \in (0,1]$$

Thus

$$\hat{x}_{\text{CE}}(y) = \mathbf{E}[X|Y=y] = \int_0^1 x \cdot 2x\,dx = 2/3.$$

(c) We first note that the conditional probability does not depend on y. Therefore, X and Y are independent, and whether or not we observe $Y = y$ does not affect the estimate in part (b). Another way to see this is to consider that if we do not observe y, we can compute the marginal $f_X(x) = \int_0^1 4xydy = 2x$ which is equal to the conditional density, and will therefore produce the same estimate.

(d) Since X and Y are independent, no estimator can make use of the observed value of Y to estimate X. The MAP estimator for X is equal to 1, regardless of what value y we observe, since the conditional (and the marginal) density is maximized at 1.

3. (a) Since the join distribution is less stretched in the Y direction, roughly speaking, knowledge of X provides more information about Y (more accurately the range of Y) than vice versa. Hence, we would choose to estimate Y based on observations of X.

To answer the rest of the questions, let us first compute the marginal distributions of X and Y :

$$f_X(x) = \int_0^{2-\frac{1}{2}x} f_{X,Y}(x,y)\,dy = \frac{1}{2} - \frac{1}{8}x, \qquad x \in [0,4]$$
$$f_Y(y) = \int_0^{4-2y} f_{X,Y}(x,y)\,dx = 1 - \frac{1}{2}y, \qquad y \in [0,2]$$

We now compute the conditional PDFs:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} 1/(4-2y) & \text{if} \quad y \in [0,2),\ x \in [0,4-2y] \\ 0 & \text{if} \quad y \in [0,2),\ x \notin [0,4-2y] \\ \text{undefined} & \text{if} \quad y < 0 \text{ or } y \geq 2 \end{cases}$$
$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \begin{cases} 2/(4-x) & \text{if} \quad x \in [0,4),\ y \in [0,2-0.5x] \\ 0 & \text{if} \quad x \in [0,4),\ y \notin [0,2-0.5x] \\ \text{undefined} & \text{if} \quad x < 0 \text{ or } x \geq 4 \end{cases}$$

(b)

$$\hat{x} = \mathbf{E}[X|Y=y] = \int x f_{X|Y}(x|y)\,dx = \frac{1}{4-2y}\int_0^{4-2y} x\,dx = 2-y, \qquad y \in [0,2)$$

(c)

$$\hat{y} = \mathbf{E}[Y|X=x] = \int y f_{Y|X}(y|x)\,dx = \frac{2}{4-x}\int_0^{2-0.5x} y\,dy = 1 - \frac{1}{4}x, \qquad x \in [0,4)$$

(d)

$$\begin{aligned} \mathbf{E}\left[(\hat{x}-X)^2 \,|Y=y\right] &= \mathbf{E}\left[\hat{x}^2 + X^2 - 2\hat{x}X \mid Y=y\right] \\ &= \frac{4}{3}y^2 - \frac{16}{3}y + \frac{16}{3} - (2-y)^2 \\ &= \frac{1}{3}(y-2)^2 \end{aligned}$$

$$\begin{aligned} \mathbf{E}\left[(\hat{y}-Y)^2 \,|X=x\right] &= \mathbf{E}\left[\hat{y}^2 + Y^2 - 2\hat{y}Y \mid X=x\right] \\ &= \frac{1}{12}x^2 - \frac{2}{3}x + \frac{4}{3} - \left(1 - \frac{1}{4}x\right)^2 \\ &= \frac{1}{3}\left(\frac{1}{4}x - 1\right)^2 \end{aligned}$$

The LMS estimate of X based on Y has worst case error of 4/3 which is realized for $Y = 0$.
The LMS estimate of Y based on X has worst case error of 1/3 which is realized for $X = 0$.

(e) Let $g(Y) = \mathbf{E}\left[(\hat{x} - X)^2 \,|Y\right]$.

$$\begin{aligned}
\mathbf{E}[g(Y)] &= \mathbf{E}[\frac{1}{3}(Y-2)^2] \\
&= \frac{1}{3}\mathbf{E}[Y^2 - 4Y + 4] \\
&= \frac{1}{3}\int_0^2 y^2(1-\frac{y}{2}) - 4y(1-\frac{y}{2})dy + \frac{4}{3} \\
&= \frac{2}{3}
\end{aligned}$$

Let $g(X) = \mathbf{E}\left[(\hat{y} - Y)^2 \,|X = x\right]$.

$$\begin{aligned}
\mathbf{E}[g(X)] &= \mathbf{E}[\frac{1}{3}(\frac{X}{4} - 1)^2] \\
&= \frac{1}{3}\mathbf{E}[\frac{1}{16}X^2 - \frac{1}{2}X + 1] \\
&= \frac{1}{3}\int_0^4 \frac{1}{16}x^2(\frac{1}{2} - \frac{y}{8}) - \frac{x}{2}(\frac{1}{2} - \frac{y}{8})dy + \frac{1}{3} \\
&= \frac{1}{18} - \frac{4}{18} + \frac{1}{3} \\
&= \frac{3}{18}
\end{aligned}$$

(f) Since the joint is constant, the MAP rule does not give meaningful results.

4. (a) Using the total probability theorem, we have

$$p_{T_1}(t) = \int_0^1 p_{T_1|Q}(t,q) f_Q(q) dq = \int_0^1 (1-q)^{t-1} q dq = \frac{1}{(t+1)t} \qquad \text{for } t = 1, 2, \ldots$$

(b) The least squares estimate coincides with the conditional expectation of Q given T_1, which is derived as

$$\begin{aligned}
\mathbf{E}[Q \mid T_1 = t] &= \int_0^1 p_{Q|T_1}(q \mid t) q dq \\
&= \int_0^1 \frac{p_{T_1|Q}(t \mid q) f_Q(q)}{p_{T_1}(t)} q dq \\
&= \int_0^1 t(t+1) q (1-q)^{t-1} q dq \\
&= \int_0^1 t(t+1) q^2 (1-q)^{t-1} dq \\
&= t(t+1)\frac{2(t-1)!}{(t+2)!} \\
&= \frac{2}{t+2}
\end{aligned}$$

(c) We write the posterior probability distribution of Q given $T_1 = t_1, \dots, T_k = t_k$

$$\begin{aligned} f_{Q|T_1,\dots,T_k}(q \mid t_1, \dots, t_k) &= \frac{f_Q(q) \prod_i^k P_{T_i}(T_i = t_i \mid Q = q)}{\int_0^1 f_Q(q) \prod_i^k P_{T_i}(T_i = t_i \mid Q = q) dq} \\ &= \frac{q^k (1-q)^{\sum_i^k t_i - k}}{c} \\ &= \frac{1}{c} q^k (1-q)^{\sum_i^k t_i - k}, \end{aligned}$$

where the denominator integrates out q so it could be viewed as a constant scalar c. To maximize the above probability we set its derivative with respect to q to zero

$$kq^{k-1}(1-q)^{\sum_i^k t_i - k} - (\sum_i^k t_i - k) q^k (1-q)^{\sum_i^k t_i - k - 1} = 0,$$

or equivalently

$$k(1-q) - (\sum_i^k t_i - k) q = 0,$$

which yields the MAP estimate

$$\hat{q} = \frac{k}{\sum_{i=1}^k t_i}.$$

For this part only assume q is sampled from the random variable Q which is now uniformly distributed over $[0.5, 1]$

(d) The LLSE of T_1 given T_2 is

$$\hat{T_2} = \mathbf{E}[T_2] + \frac{\text{cov}(T_1, T_2)}{\text{var}(T_1)} (T_1 - \mathbf{E}[T_1]),$$

where the coefficients are

$$\mathbf{E}[T_1] = \mathbf{E}[T_2] = \int_{0.5}^1 f_Q(q) \mathbf{E}[T|Q = q] dq = \int_{0.5}^1 2 * 1/q dq = 2 \ln 2,$$

and from the law of total variance

$$\begin{aligned} \text{var}(T_1) = \text{var}(T_2) &= \mathbf{E}\left[\text{var}(T_1 \mid Q)\right] + \text{var}\left[\mathbf{E}(T_1 \mid Q)\right] \\ &= \mathbf{E}\left[\frac{1-Q}{Q^2}\right] + \text{var}\left[\frac{1}{Q}\right] \\ &= \mathbf{E}[1/Q^2] - \mathbf{E}[1/Q] + \mathbf{E}[1/Q^2] - \mathbf{E}[1/Q]^2 \\ &= \int_{0.5}^2 f_Q(q) \frac{1}{q^2} dq - \int_{0.5}^2 f_Q(q) \frac{1}{q} dq + \int_{0.5}^2 f_Q(q) \frac{1}{q^2} dq - \left(\int_{0.5}^2 f_Q(q) \frac{1}{q} dq \right)^2 \\ &= 2 - 2 \ln 2 + 2 - (2 \ln 2)^2 \\ &= 4 - 2 \ln 2 - (2 \ln 2)^2, \end{aligned}$$

and their covariance

$$\begin{aligned}\mathrm{cov}(T_1, T_2) &= \mathbf{E}[T_1T_2] - \mathbf{E}[T_1]\mathbf{E}[T_2] \\ &= \mathbf{E}\left[\mathbf{E}[T_1T_2 \mid Q]\right] - \mathbf{E}[T_1]\mathbf{E}[T_2] \\ &= \mathbf{E}\left[\mathbf{E}[T_1 \mid Q]\mathbf{E}[T_2 \mid Q]\right] - \mathbf{E}[T_1]\mathbf{E}[T_2] \\ &= \mathbf{E}\left[1/Q^2\right] - \mathbf{E}[T_1]\mathbf{E}[T_2] \\ &= 2 - 4(\ln 2)^2\end{aligned}$$

Therefore we have derived the linear least squares estimator

$$\hat{T}_2 = 2\ln 2 + \frac{2 - 4(\ln 2)^2}{4 - 2\ln 2 - (2\ln 2)^2}(T_1 - 2\ln 2) \approx 1.543 + 0.113T_1.$$

5. (a)

$$\mathrm{cov}(X, Y) \;=\; \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y]$$

With independence required in some steps,

$$\begin{aligned}\mathbf{E}[X] &= \mathbf{E}[W_1 + W_2] \;=\; \mathbf{E}[W_1] + \mathbf{E}[W_2] \;=\; \frac{1}{2} + \frac{1}{2} \;=\; 1 \\ \mathbf{E}[Y] &= \mathbf{E}[X + W_2] \;=\; \mathbf{E}[X] + \mathbf{E}[W_2] \;=\; 1 + \frac{1}{2} \;=\; \frac{3}{2} \\ \mathrm{var}(X) &= \mathrm{var}(W_1 + W_2) \;=\; \mathrm{var}(W_1) + \mathrm{var}(W_2) \;=\; \frac{1}{12} + \frac{1}{12} \;=\; \frac{1}{6} \\ \mathbf{E}[XY] &= \mathbf{E}[X(X + W_3)] \;=\; \mathbf{E}[X^2] + \mathbf{E}[XW_3] \;=\; \mathrm{var}(X) + (\mathbf{E}[X])^2 + \mathbf{E}[X]\mathbf{E}[W_3] \\ &= \frac{1}{6} + 1^2 + 1 \cdot \frac{1}{2} \;=\; \frac{5}{3} \\ \mathrm{cov}(X, Y) &= \mathbf{E}[XY] - \mathbf{E}[X]\,\mathbf{E}[Y] \;=\; \frac{5}{3} - 1 \cdot \frac{3}{2} \;=\; \frac{1}{6}.\end{aligned}$$

(b) The linear least mean squares estimator of X from Y is given by

$$\widehat{X}_{LLMS} \;=\; \frac{\mathrm{cov}(X, Y)}{\mathrm{var}(Y)}(Y - \mathbf{E}[Y]) + \mathbf{E}[X],$$

and the required quantities are straightforward to compute. With independence required in some steps and using the results from part (a),

$$\mathrm{var}(Y) \;=\; \mathrm{var}(W_1 + W_2 + W_3) \;=\; \mathrm{var}(W_1) + \mathrm{var}(W_2) + \mathrm{var}(W_3) \;=\; \frac{1}{12} + \frac{1}{12} + \frac{1}{12} \;=\; \frac{1}{4}.$$

Thus, the desired estimator is

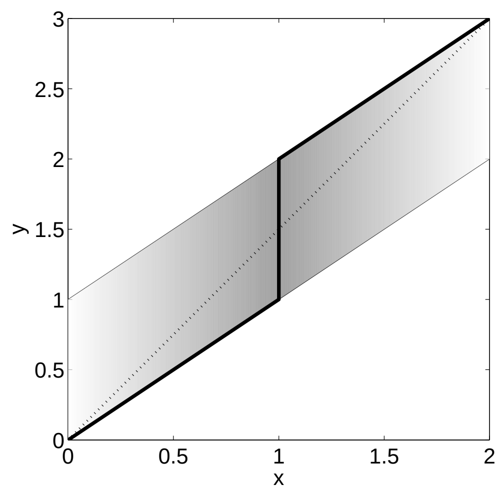
$$\widehat{X}_{LLMS} \;=\; \frac{2}{3}\left(Y - \frac{3}{2}\right) + 1 \;=\; \frac{2}{3}Y.$$

It is also possible to find the LMS estimator (without presuming it to be linear) and then notice that it is linear.

(c) As the sum of two uniform random variables, X has the triangular PDF

$$f_X(x) \;=\; \begin{cases} 1-|x-1|, & \text{for } x \in [0,2]; \\ 0, & \text{otherwise.} \end{cases}$$

The conditional PDF $f_{Y|X}(y \mid x)$ is uniform over $[x, x+1]$. Thus the joint PDF $f_{X,Y}(x,y)$ is nonzero on a parallelogram as marked below, with constant value on vertical slices within the parallelogram but non-constant value on horizontal slices. This is depicted with shading below.



The MAP estimate of X from $Y = y$ is obtained by finding the maximum of $f_{X,Y}(x,y)$ along the horizontal slice determined by $Y = y$. This maximum is obtained on the bold curve above. Thus,

$$\widehat{X}_{MAP} \;=\; \begin{cases} Y, & \text{for } Y \in [0,1); \\ 1, & \text{for } Y \in [1,2]; \\ Y-1, & \text{for } Y \in (2,3]. \end{cases}$$

(The LLMS estimator is shown with a dotted line.)

G1†. (a) For convenience of notation, lets say that all random variables X_k have the same distribution as a random variable X. Since

$$\log_b(R_n) = \frac{1}{n}\sum_{k=1}^{n} \log_b(X_k),$$

an average, where the terms $\log_b(X_k)$ are independent and identically distributed (since the X_ks are), and since we assume the mean and variance of $\log_b(X_k)$ exist and are finite, it follows from the weak law of probability that the sequence $\frac{1}{n}\sum_{k=1}^{n} \log_b(X_k)$ converges in probability to the expectation of $\log_b(X_k)$, i.e., prob.

$$\log_b(R_n) \xrightarrow[n\to\infty]{prob.} \mathbf{E}[\log_b(X)].$$

(b) Since $R_n = b^{\log_b(R_n)}$, where the function $x \to b^x$ is continuous, it follows from the lemma and the solution to part (a) that

$$R_n \xrightarrow[n\to\infty]{prob.} r = b^{\mathbf{E}[\log_b(X)]}.$$

For the Double or Quarter game with $p = 1/2$, letting $b = 2$, we have:

$$\mathbf{P}(X = 2) = \mathbf{P}(X = 1/4) = 1/2$$
$$\mathbf{P}(\log_2 X = 1) = \mathbf{P}(\log_2 X = -2) = 1/2$$
$$\mathbf{E}[\log_2(X)] = (1/2)(1) + (1/2)(-2) = -1/2,$$

from which we conclude that

$$R_n \xrightarrow[n\to\infty]{prob.} r = 2^{-1/2} = \frac{1}{\sqrt{2}} \approx 0.7071.$$

This corresponds to an average loss of approximately 29% per game!

(c) Since $(R_n)^n = W_n = X_1 \cdot X_2 \cdot \cdots \cdot X_n$ and $R_n \xrightarrow[n\to\infty]{prob.} r \approx 0.7071 < 1$, it follows that $R_n^n = W_n \xrightarrow[n\to\infty]{prob.} 0$.

To see this a bit more rigorously, note that since R_n is a positive quantity and R_n converges in probability to r, $0 < r < 1$, if we pick a point between r and 1, say $\frac{r+1}{2}$ (where $r < \frac{r+1}{2} < 1$), that $\mathbf{P}\left(R_n \geq \frac{r+1}{2}\right) \xrightarrow[n\to\infty]{prob.} 0$. But since the event $\left\{R_n \geq \frac{r+1}{2}\right\}$ is the same as the event $\left\{(R_n)^n \geq \left(\frac{r+1}{2}\right)^n\right\}$ and by definition, $W_n\, (R_n)^n$, it follows that

$$\mathbf{P}\left(R_n \geq \frac{r+1}{2}\right) = \mathbf{P}\left((R_n)^n \geq \left(\frac{r+1}{2}\right)^n\right) = \mathbf{P}\left(W_n \geq \left(\frac{r+1}{2}\right)^n\right),$$

and since $0 < \frac{r+1}{2} < 1$, it also follows $\left(\frac{r+1}{2}\right)^n \xrightarrow[n\to\infty]{} 0$, which implies that for any $\epsilon > 0$, $\mathbf{P}(W_n \geq \epsilon) = \mathbf{P}(|W_n - 0| \geq \epsilon) \xrightarrow[n\to\infty]{} 0$.

(d) A picture really is helpful here.

Since the logarithm function is concave (i.e., $\frac{d\log(x)}{dx} < 0$ everywhere), the tangent line to $y = \log(x)$ at any point lies strictly above the plot of $y = \log(x)$. In particular, the tangent line at $x = \mathbf{E}[X]$ lies above the plot and is given by the equation:

$$y = f_L(x) = \log(\mathbf{E}[X]) + \frac{d\log(x)}{dx}\bigg|_{x=\mathbf{E}[X]}(x - \mathbf{E}[X]).$$

Since $f_L(x) \geq \log(x)$ for all $x > 0$, it follows that

$$\mathbf{E}[f_L(X)] = \log(\mathbf{E}[X]) \geq \mathbf{E}[\log(X)].$$

.

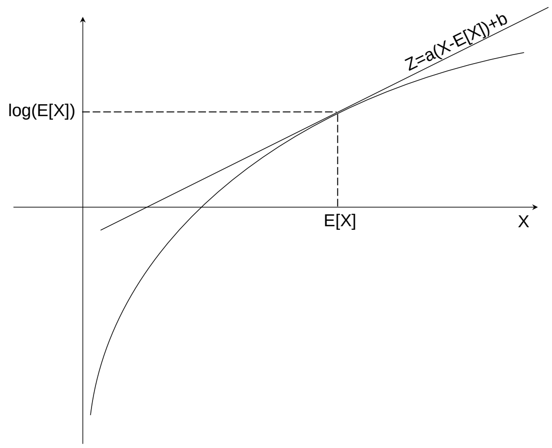


Figure 1: The log function

Therefore, since b^y is monotone increasing in y for any $b > 1$,

$$R_n = b^{\frac{1}{n}\sum_{k=1}^{n}\log_b(X_k)} \xrightarrow[n\to\infty]{prob.} r = b^{\mathbf{E}[\log_b(X)]} \leq b^{\log_b(\mathbf{E}[X])} = \mathbf{E}[X_k]$$

The only case where $r = \mathbf{E}[X]$ is the case where $\mathbf{P}(X = c > 0) = 1$.

(e) The conclusion to part (b), $R_n = \log_b(X_k) \xrightarrow[n\to\infty]{prob.} r = b^{\mathbf{E}[\log_b(X)]}$ shows us that, if $r > 1$, your wealth grows toward $+\infty$ as $n \to \infty$, and if $r < 1$, your wealth shrinks toward zero as $n \to \infty$.

Using the fixed fraction strategy, where X_n is replaced by the function of f: We are given

$$X_n(f) = [(1-f) + fX_n].$$

(This way of gambling, which takes the "edge" off the game, has some surprising consequences, as we will see.)

$$\begin{aligned}
\mathbf{E}[\log_2((1-f)+fX)] &= (1/2)\log_2(1-f+(1/4)f) + (1/2)\log_2(1-f+2f) \\
&= (1/2)\log_2[(1-3f/4)(1+f)] \\
&= (1/2)\log_2[(1+f/4-3f^2/4)].
\end{aligned}$$

To find the maximum, we differentiate:

$$\frac{\partial}{\partial f}[1+f/4-3f^2/4] = [1/4-(6/4)f] = 0.$$

i.e. $f = 1/6$. We know this is the maximum since the log is monotonically increasing and

$$\frac{\partial^2}{\partial f^2}[1+f/4-3f^2/4] = -6/4 < 0.$$

The function of f, $1 + f/4 - 3f^2/4$, equals 1 at $f = 0$ and $f = 1/3$ and is positive between those values and negative elsewhere. Therefore, returning to the original statement of the problem,

$$R_n \xrightarrow[n\to\infty]{prob.} r = 2^{\mathbf{E}[\log_2 X(f)]},$$

we see that

$$\mathbf{E}[\log_2(X(f))] = (1/2)\log_2[(1 + f/4 - 3f^2/4)],$$

where

$$\log_2[(1 + f/4 - 3f^2/4)] > 0 \Leftrightarrow 0 < f < 1/3,$$

so that

$$r > 1 \Leftrightarrow 0 < f < 1/3,$$

which implies that this your financial future if you play this game: your wealth W_n

$$W_n \xrightarrow[n\to\infty]{prob.} \infty \text{ if } 0 < f < 1/3,$$

$$W_n \xrightarrow[n\to\infty]{prob.} 0 \text{ if } 1/3 < f \leq 1.$$

The maximum value of $\mathbf{E}[\log_2(X(f))]$ occurs at $f = 1/6$, a value at which

$$r = 2^{\mathbf{E}[\log_2 X(f)]} = 2^{(1/2)\log_2[(1+f/4-3f^2/4)]}\Big|_{f=1/6} \approx 2^{(1/2)\log_2[1.02083]} \approx 1.0104,$$

i.e., your wealth will grow, on average, about 1.04% per toss!!