(d) Show that
\[ \Pr\{V_1 > V_0 \mid X=0, \phi=0\} = \int f_{Y_1,Y_2}(y_1,y_2 \mid 0,0) \Pr\left\{ V_1 > y_1^2 + y_2^2 \right\} dy_1 dy_2. \]
Show that this is equal to \((1/2) \exp(-a^2/(4\sigma^2)).\)

(e) Explain why this is the probability of error (i.e., why the event \(V_1 > V_0\) is independent of \(\phi\)), and why \(\Pr\{e \mid X=0\} = \Pr\{e \mid X=1\}.

Exercise 8.8 Binary FSK on a Rayleigh fading channel can be modeled in terms of a four-dimensional observation vector \(Y = (Y_1, Y_2, Y_3, Y_4)^T\), where \(Y = U + Z\). The L-rv \(Z \sim N(0, \sigma^2[I])\) is independent of \(U\). Under \(X = 0\), \(U = (U_1, U_2, 0, 0)^T\), whereas under \(X = 1\), \(U = (0, 0, U_3, U_4)^T\). The rvs \(U_i \sim N(0, \alpha^2)\) are IID. The a priori probabilities are \(p_0 = p_1 = 1/2\).

(a) Convince yourself from the circular symmetry of the situation that the ML receiver calculates sample values \(v_0\) and \(v_1\) for \(V_0 = Y_1^2 + Y_2^2\) and \(V_1 = Y_3^2 + Y_4^2\) and chooses \(\hat{x} = 0\) if \(v_0 > v_1\) and chooses \(\hat{x} = 1\) otherwise.

(b) Find \(f_{v_0|X}(v_0 \mid 0)\) and find \(f_{v_1|X}(v_1 \mid 0).\)

(c) Let \(W = V_0 - V_1\) and find \(f_w(w \mid X=0).\)

(d) Show that \(\Pr\{e \mid X=0\} = [2 + a^2/\sigma^2]^2 - 1.\) Explain why this is also the unconditional probability of an incorrect decision.

Exercise 8.9 A disease has two strains, 0 and 1, which occur with a priori probabilities \(p_0\) and \(p_1 = 1 - p_0\) respectively.

(a) Initially, a rather noisy test was developed to find which strain is present for patients with the disease. The output of the test is the sample value \(y_1\) of a rv \(Y_1\). Given strain 0 (\(X = 0\), \(Y_1 = 5 + Z_1\), and given strain 1 (\(X = 1\), \(Y_1 = 1 + Z_1\). The measurement noise \(Z_1\) is independent of \(X\) and is Gaussian, \(Z_1 \sim N(0, \sigma^2)\). Give the MAP decision rule, i.e., determine the set of observations \(y_1\) for which the decision is \(\hat{x} = 1).\) Give \(\Pr\{e \mid X=0\}\) and \(\Pr\{e \mid X=1\}\) in terms of the function \(Q(x).\)

(b) A budding medical researcher determines that the test is making too many errors. A new measurement procedure is devised with two observation rvs \(Y_1\) and \(Y_2\). \(Y_1\) is the same as in (a). \(Y_2,\) under hypothesis 0, is given by \(Y_2 = 5 + Z_1 + Z_2,\) and, under hypothesis 1, is given by \(Y_2 = 1 + Z_1 + Z_2\). Assume that \(Z_2\) is independent of both \(Z_1\) and \(X,\) and that \(Z_2 \sim N(0, \sigma^2).\) Find the MAP decision rule for \(\hat{x}\) in terms of the joint observation \((y_1,y_2)\), and find \(\Pr\{e \mid X=0\}\) and \(\Pr\{e \mid X=1\}.\) Hint: Find \(f_{y_2|y_1,X}(y_2 \mid y_1,0)\) and \(f_{y_2|y_1,X}(y_2 \mid y_1,1).\)

(c) Explain in laymen’s terms why the medical researcher should learn more about probability.

(d) Now suppose that \(Z_2,\) in (b), is uniformly distributed between 0 and 1 rather than being Gaussian. We are still given that \(Z_2\) is independent of both \(Z_1\) and \(X.\) Find the MAP decision rule for \(\hat{x}\) in terms of the joint observation \((y_1,y_2)\) and find \(\Pr\{e \mid X=0\}\) and \(\Pr\{e \mid X=1\).\)

(e) Finally, suppose that \(Z_1\) is also uniformly distributed between 0 and 1. Again find the MAP decision rule and error probabilities.
Exercise 8.10  (a) Consider a binary hypothesis testing problem, and denote the hypotheses as $X = 1$ and $X = -1$. Let $a = (a_1, a_2,\ldots, a_n)^T$ be an arbitrary real $n$-vector and let the observation be a sample value $y$ of the rv $Y = Xa + Z$, where $Z \sim \mathcal{N}(0, \sigma^2[I_n])$ and $[I_n]$ is the $n \times n$ identity matrix. Assume that $Z$ and $X$ are independent. Find the ML decision rule and find the probabilities of error $\Pr(e | X = 0)$ and $\Pr(e | X = 1)$ in terms of the function $Q(x)$.

(b) Now suppose a third hypothesis, $X = 0$, is added to the situation of (a). Again the observation rv is $Y = Xa + Z$, but here $X$ can take on values $-1, 0, \text{ or } +1$. Find a one-dimensional sufficient statistic for this problem (i.e., a one-dimensional function of $y$ from which the likelihood ratios \[ \Lambda_1(y) = \frac{p_{Y|X}(y | 1)}{p_{Y|X}(y | 0)} \quad \text{and} \quad \Lambda_{-1}(y) = \frac{p_{Y|X}(y | -1)}{p_{Y|X}(y | 0)} \] can be calculated).

(c) Find the ML decision rule for the situation in (b) and find the probabilities of error, $\Pr(e | X = x)$ for $x = -1, 0, +1$.

(d) Now suppose that $Z_1,\ldots,Z_n$ in (a) are IID and each is uniformly distributed over the interval $-2$ to $+2$. Also assume that $a = (1, 1,\ldots,1)^T$. Find the ML decision rule for this situation.

Exercise 8.11  A sales executive hears that one of his salespeople is routing half of his incoming sales to a competitor. In particular, arriving sales are known to be Poisson at rate one per hour. According to the report (which we view as hypothesis $X = 1$), each second arrival is routed to the competition; thus under hypothesis 1 the interarrival density for successful sales is $f(y | X = 1) = ye^{-y}; y \geq 0$. The alternative hypothesis ($X = 0$) is that the rumor is false and the interarrival density for successful sales is $f(y | X = 0) = e^{-y}; y \geq 0$. Assume that, a priori, the hypotheses are equally likely. The executive, a recent student of stochastic processes, explores various alternatives for choosing between the hypotheses; he can only observe the times of successful sales, however.

(a) Starting with a successful sale at time 0, let $S_i$ be the arrival time of the $i$th subsequent successful sale. The executive observes $S_1, S_2,\ldots,S_n (n \geq 1)$ and chooses the maximum a posteriori probability hypothesis given this observation. Find the joint probability density $f(S_1, S_2,\ldots,S_n | X = 1)$ and $f(S_1,\ldots,S_n | X = 0)$ and give the decision rule.

(b) This is the same as (a) except that the system is in steady state at time 0 (rather than starting with a successful sale). Find the density of $S_1$ (the time of the first arrival after time 0) conditional on $X = 0$ and on $X = 1$. What is the decision rule now after observing $S_1,\ldots,S_n$.

(c) This is the same as (b), except rather than observing $n$ successful sales, the successful sales up to some given time $t$ are observed. Find the probability, under each hypothesis, that the first successful sale occurs in $(s_1, s_1 + \Delta]$, the second in $(s_2, s_2 + \Delta],\ldots,$ and the last in $(s_{N(t)}, s_{N(t)} + \Delta]$ (assume $\Delta$ very small). What is the decision rule now?
Exercise 8.12  This exercise generalizes the MAP rule to cases where neither a PMF
nor a PDF exists. To view the problem in its simplest context, assume a binary decision,
a one-dimensional observation ($y \in \mathbb{R}^1$), and fixed a priori probabilities $p_0, p_1$. An arbitrary
rule, denoted test $A$, can be defined by the set $A$ of observations that are mapped
into decision $\hat{x} = 1$. Using such a test, the overall probability of correct decision is given
by (8.6) as

$$\Pr\{\hat{X}_A(Y) = X\} = p_0 \Pr\{Y \in A^c \mid X = 0\} + p_1 \Pr\{Y \in A \mid X = 1\}.$$  (8.84)

The maximum of this probability over all tests $A$ is then

$$\sup_A \Pr\{\hat{X}_A(Y) = X\} = \sup_A \left[ p_0 \Pr\{Y \in A^c \mid X = 0\} + p_1 \Pr\{Y \in A \mid X = 1\}\right].$$  (8.85)

We first consider this supremum over all $A$ consisting of a finite union of disjoint inter-
vals. Then (using measure theory), we show that the supremum over all measurable sets
$A$ is the same as that over finite unions of disjoint intervals.

(a) If $A$ is a union of $k$ intervals, show that $A^c$ is a union of at most $k + 1$ disjoint
intervals. Intervals can be open or closed on each end and can be bounded by $\pm \infty$.

(b) Let $I$ be the partition of $\mathbb{R}$ created by the intervals of both $A$ and of $A^c$. Let $I_j$ be
the $j$th interval in this partition. Show that

$$\Pr\{\hat{x}_A(Y) = X\} \leq \sum_j \max\left[p_0 \Pr\{Y \in I_j \mid X = 0\}, p_1 \Pr\{Y \in I_j \mid X = 1\}\right].$$

Hint: Break (8.84) into intervals and apply the MAP principle on an interval basis.

(c) The expression on the right in (b) is a function of the partition but is otherwise
independent of $A$. It corresponds to a test where $Y$ is first quantized to a finite set of
intervals, and the MAP test is then applied to this discrete problem. We denote this test
as MAP($I$). Let the partition $I'$ be a refinement of $I$ in the sense that each interval of $I'$
is contained in some interval of $I$. Show that

$$\Pr\{\hat{x}_{\text{MAP}(I')} (Y) = X\} \leq \Pr\{\hat{x}_{\text{MAP}(I)}(Y) = X\}.$$  (d)

(d) Show that for any two finite partitions of $\mathbb{R}$ into intervals, there is a third finite
partition that is a refinement of each of them.

(e) Consider a sequence $\{A_j, j \geq 1\}$ (each is a finite union of intervals) that approaches
the supremum (over finite unions of intervals) of (8.85). Demonstrate a corresponding
sequence of successive refinements of partitions $\{I_j, j \geq 1\}$ for which $\Pr(\hat{x}_{\text{MAP}(I_j)}(Y) = X)$ approaches the same limit.

Note what this exercise has shown so far: if the likelihoods have no PDF or PMF,
there is no basis for a MAP test on an individual observation $y$. However, by quantizing
the observations sufficiently finely, the quantized MAP rule has an overall probability of
being correct that is as close as desired to the optimum over all rules where $A$ is a finite
(but arbitrarily large) union of intervals. The arbitrarily fine quantization means that the
decisions are arbitrarily close to pointwise decisions, and the use of quantization means
that infinitely fine resolution is not required for the observations.
(f) Next suppose the supremum in (8.85) is taken over all measurable sets $A \in \mathbb{R}$. Show that, given any measurable set $A$ and any $\epsilon > 0$, there is a finite union of intervals $A'$ such that $\Pr \{ A \neq A' \mid X = \ell \} \leq \epsilon$ both for $\ell = 0$ and $\ell = 1$. Show from this that the supremum in (8.85) is the same whether taken over measurable sets or finite unions of intervals.

Exercise 8.13 For the minimum-cost hypothesis testing problem of Section 8.3, assume that there is a cost $C_0$ of choosing $X = 1$ when $X = 0$ is correct, and a cost $C_1$ of choosing $X = 0$ when $X = 1$ is correct. Show that a threshold test minimizes the expected cost using the threshold $\eta = (C_0 p_0)/(C_1 p_1)$.

Exercise 8.14 (a) For given $\theta$, $0 < \theta \leq 1$, let $\eta^*$ achieve the supremum $\sup_{0 \leq \eta < \infty} q_1(\eta) + \eta (q_0(\eta) - \theta)$. Show that $\eta^* \leq 1/\theta$. Hint: Think in terms of Lemma 8.4.1 applied to a very simple test.

(b) Show that the magnitude of the slope of the error curve $u(\theta)$ at $\theta$ is at most $1/\theta$.

Exercise 8.15 Consider a binary hypothesis testing problem where $X$ is 0 or 1 and a one-dimensional observation $Y$ is given by $Y = X + U$, where $U$ is uniformly distributed over $[-1, 1]$ and is independent of $X$.

(a) Find $f_{Y|X}(y \mid 0)$, $f_{Y|X}(y \mid 1)$ and the likelihood ratio $\Lambda(y)$.

(b) Find the threshold test at $\eta$ for each $\eta$, $0 < \eta < \infty$ and evaluate the conditional error probabilities, $q_0(\eta)$ and $q_1(\eta)$.

(c) Find the error curve $u(\theta)$ and explain carefully how $u(0)$ and $u(1/2)$ are found. Hint: $u(0) = 1/2$.

(d) Find a discrete sufficient statistic $v(y)$ for this problem that has three sample values.

(e) Describe a decision rule for which the error probability under each hypothesis is $1/4$. You need not use a randomized rule, but you need to handle the don’t-care cases under the threshold test carefully.