General probability questions

1. Exercises 2E1 in the book.
2. Exercises 2E2 in the book.
3. Exercises 2E3 in the book.

Bayes theorem

1. In this exercise we will use Bayes theorem for spell correction. Let θ be the word one wanted to spell, and let Y be the word that is actually spelled, which in this case is $'radom'$. Suppose we know that it should be one of the three words $\{'radom','random',radon'\}$. The three words are not equally like a priori (unnormalized prior taken from frequency of the word occurring):

θ	$P(\theta)$
$'random'$	7×10^{-5}
$'radon'$	7×10^{-6}
$'radom'$	3×10^{-7}

The conditional probability of θ given one has spelled $'radom'$ (the likelihood) is, by Google spell checker, given by (again not normalized):

θ	$P(Y =' radom'\|\theta)$
$'random'$	0.00193
$'radon'$	0.000143
$'radom'$	0.975

What is the posterior distribution of θ given that $Y =' radom'$?

Proportionality

Recall that two functions $f_1(x)$ and $f_2(x)$ are proportional, $\propto$, to each other if there exists a positive constant $c > 0$ such that $f_1(x) = cf_2(x)$.

1. Suppose $f(x) = e^{-x^2}$ which of the following functions are $f(x)$ proportional to
   a) $f_1(x) = e^{-\frac{x^2}{2}}$.
   b) $f_2(x) = e^{-(x-\mu)^2}$.
   c) $f_3(x) = e^{-x^2-\mu^2}$.
2. Suppose $f(p) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}$ which of the following functions are $f(p)$ proportional to
   a) $f_1(x) = x^{\alpha-1}(1-x)^{\beta-1}$.
   b) $f_2(x) = x^{\alpha+2-1}(1-x)^{\beta-1}$.
   c) $f_3(x) = x^{\alpha-1}(1-x)^{\beta+3-1}$.

Independence

Independence recall that the definition of independence for probabilities is given by

$$P(A \cap B) = P(A)P(B).$$

The same holds for densities, i.e.

$$p_{X,Y}(x, y) = p_X(x)p_Y(y).$$

1. Suppose that the joint density of X_1 and X_2 given p is proportional to:

$$p(x_1, x_2|p) \propto p^{x_1+x_2}(1-p)^{2-x_1-x_2}$$

   Show that given p that X_1 and X_2 two are independent.

2. Suppose we have a uniform prior on p (i.e. $p(p) = 1$), and suppose we observe n independent observations of a Bernoulli random variable (i.e. $p(X|p) = p^X(1-p)^{1-X}$). Show that (using Bayes theorem and independence)

$$p(p|X_1, X_2, \ldots, X_n) \propto p^{\sum_{i=1}^n X_i}(1-p)^{n-\sum_{i=1}^n X_i}$$

Distributions

A key concept in Bayesian modeling is distribution.

1. The most famous distribution is the Normal distribution:

$$p(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{(x-\mu)^2}{2\sigma^2}}.$$

   a) Is the distribution of x continuous or discrete?
   b) Suppose we want to put a prior on μ, if we have no special information should the prior be continuous or discrete?
   c) Suppose we want to put a prior on σ, if we have no special information should the prior be continuous or discrete? Should there be any other restriction on the domain of σ (which values σ can take)?
   d) Write the integral form of $P(X > 2|\mu = 2, \sigma = 4)$.

2. An other classical distribution is the Poisson distribution:

$$p(x|\lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \quad x = 0, 1, 2 \ldots$$

   a) Is the distribution of x continuous or discrete?
   b) Suppose we want to put a prior on λ, if we have no special information should the prior be continuous or discrete?
   c) Write the sum form of $P(X > 2|\lambda = 2)$
   d) Suppose we observe two independent observation (X_1, X_2) observations of $p(x|\lambda)$. Further assume a an exponential prior on λ ($p(\lambda) = e^{-\beta\lambda}$). Use Bayes theorem and independence to derive the posterior distribution of λ ($p(\lambda|X_1, X_2)$)