Questions

1. (a) Suppose we have a normally (Gaussian) distributed random variable with mean \( \mu \) and variance \( \sigma^2 \). Suppose we sample \( n \) times from this distribution. What is the likelihood function \( p(\vec{x} \mid \mu, \sigma^2) \) where \( \vec{x} \) are the samples?

(b) What is the value of \( \mu \) that gives the maximum of this likelihood function?

2. Suppose you have a random variable \( X \) that takes value 0 with probably \( p_0 \) and 1 with probability \( 1 - p_0 \). Assume \( 0 < p_0 < 1 \).

(a) Suppose you sample this random variable repeatedly until you get a value 1 and it takes you \( k \) tries, i.e. a 1 first occurs on trial \( k \). What is the likelihood function \( p(k \mid p_0) \) ?

(b) What is the likelihood for \( p_0 = 0 \) or \( 1 \)? What is the limit of the likelihood function as \( p_0 \) goes to 0 or 1?

(c) For \( k = 1 \), what is \( \lim_{p_0 \to 0} p(k = 1 \mid p_0) \) ?

3. Suppose we have two likelihood functions \( p(I_1 \mid S) \) and \( p(I_2 \mid S) \) that we assume are ”conditionally independent”, which means:

\[
p(I_1, I_2 \mid S) = p(I_1 \mid S) p(I_2 \mid S)
\]

Intuitively, once the scene is fixed, knowing the value of one image variable \( I_1 \) tells us nothing about the value of the other image variable \( I_2 \). For example, \( I_1 \) might be the sizes of texture elements and \( I_2 \) might be the foreshortening of texture elements. Or \( I_2 \) might be the binocular disparities. Note: this is just a model. In fact, there might be a weak dependence, but we ignore this dependence to keep the model simple.

Now suppose the likelihood functions \( p(I_1 \mid S) \) and \( p(I_2 \mid S) \) are both Gaussians. Given \( I_1 = i_1 \) and \( I_2 = i_2 \), each likelihood function gives a maximum likelihood estimate which we’ll call \( S = s_1 \) and \( S = s_2 \), respectively.

What is the \( s \) that maximizes the likelihood of \( p(I_1, I_2 \mid S) \) ?
Solutions

1. (a) Assuming the samples are independent, the likelihood function is the product of individual likelihoods,
\[ p(\bar{x} | \mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}} \]
Note the variables in this likelihood function are \( \mu \) and \( \sigma \).
(b) Take the derivative with respect to \( \mu \), set it to 0, and solve for \( \mu \). This gives the maximum likelihood for \( \mu \) is just the sample mean:
\[ \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i. \]
Note that its almost never going to happen that \( \bar{x} = \mu \) since the sample mean is itself a (continuous valued) random variable.

2. (a) The likelihood function is \( p(k|p_0) = p_0^{k-1}(1-p_0) \). Note this is unimodal function of \( p_0 \).
(b) This formula only makes sense when \( 0 < p_0 < 1 \). If \( p_0 = 0 \), then \( k \) would always be 1 and so the first term would be \( 0^0 \) which is not well defined. If \( p_0 = 1 \) then the value 0 always occur and so \( k = \infty \).
For any finite \( k > 1 \) it has a limit of 0 at \( p_0 = 0 \) and at \( p_0 = 1 \).
(c) If \( k = 1 \) and \( p_0 > 0 \), then \( p(k=1|p_0) = 1 - p_0 \).

3. Multiply the two conditional probability functions, and take the derivative with respect to the scene variable \( s \) and set this derivative to 0. Note that the variances \( \sigma_1^2 \) and \( \sigma_2^2 \) need to be combined to give the variance of the likelihood \( p(I_1, I_2 | S) \).
We want to maximize
\[ p(I_1|S)p(I_2|S) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(s-s_1)^2}{2\sigma_1^2}} e^{-\frac{(s-s_2)^2}{2\sigma_2^2}} \]
so we want to minimize
\[ \frac{(s-s_1)^2}{2\sigma_1^2} + \frac{(s-s_2)^2}{2\sigma_2^2} \]
Take the derivative with respect to \( s \) and set it to 0. This gives
\[ \frac{s-s_1}{\sigma_1^2} + \frac{s-s_2}{\sigma_2^2} = 0 \]
and so
\[ s = \left( \frac{s_1}{\sigma_1^2} + \frac{s_2}{\sigma_2^2} \right) / \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right) \]
Note that this is of the form
\[ s = \sum_{i=1}^{2} w_is_i \text{ where } 0 < w_i < 1 \text{ and } w_1 + w_2 = 1. \]