Stamatics Assignment 2

Bhavya Sharma

May 2024

1 Problem Solutions

1.1 Problem 1

(a) (proof)

Let Φ be any event in the sample space Ω. Then, by definition of the preimage:

$$X^{-1}(\Phi) = \{\omega \in \Omega \mid X(\omega) \in \Phi\}$$

Since -1 is not a possible value for any random variable, $X(\omega)$ will never be in Φ. Therefore, the preimage of any event Φ under X will be the empty set, which is denoted by $\emptyset$.

We can express this mathematically:

$$X^{-1}(\Phi) = \emptyset$$

(b) (proof)

Following the same logic as part (a), the preimage of the entire sample space Ω under X will include all elements where $X(\omega)$ can take any value. Since a random variable can take any value within its support, this means:

$$X^{-1}(\Omega) = \Omega$$

1.2 Problem 2

The support of a random variable X refers to the set of all possible values for which the probability density function (pdf) is non-zero.

The distribution function (CDF) of the exponential distribution with parameter λ is given by:

$$F_X(x) = \{\, 0 \,, for x < 0$$

$$1 - e^{-\lambda x}, for x \geq 0$$

Since the CDF approaches 1 as x approaches positive infinity, the pdf $f_X(x)$ will be non-zero for all non-negative values ($x \geq 0$).

Therefore, the support of the random variable X from the exponential distribution Exp(λ) is:

1.3 Problem 3

Let's denote:

y_i = actual value of the i-th data point $\hat{y_i}$ = predicted value of the i-th data point by the linear regression model * x_i = independent variable value of the i-th data point * N = number of data points

The L2 norm loss function can be written as:

$$L = \sum_{i=1}^{N} (y_i - \hat{y_i})^2$$

We can represent the linear regression model as:

$$\hat{y_i} = mx_i + c$$

Substituting this into the loss function:

$$L = \sum_{i=1}^{N} (y_i - (mx_i + c))^2$$

To minimize L, we take partial derivatives with respect to m and c and set them equal to zero.

Deriving m:

$$\frac{\partial L}{\partial m} = -2 \sum_{i=1}^{N} (y_i - mx_i - c)x_i = 0$$

Deriving c:

$$\frac{\partial L}{\partial c} = -2 \sum_{i=1}^{N} (y_i - mx_i - c) = 0$$

Solving these equations simultaneously for m and c will lead to the expressions you provided. The specific solution steps might involve further algebraic manipulation for simplification.

Minimizing the Loss Function:

The second partial derivative test can be used to confirm that these values correspond to a minimum of the loss function. However, the derivation above already demonstrates the intuition behind minimizing the squared errors. By setting the partial derivatives to zero, we ensure that small changes in m and c won't further reduce the total error.