3F3 – Statistical Signal Processing
Examples Paper 1

Questions marked as '†' are straightforward questions testing fundamental concepts. The rest are Tripos style although not necessarily Tripos length.

1. †An incoming email is either spam or not. Let B be the event the email contains the word "free." From experience (or training data), $P(B|\text{spam}) = 0.8$ and $P(B|\text{not spam}) = 0.1$ and spam emails are 25% of all my emails. Give (Ω, P).

2. †A fair die is thrown. Let $A = \{2, 4, 6\}$ and $B = \{1, 2, 3, 4\}$. Show A and B are independent.

3. †(Inverting the cdf.)

   (a) Flip a fair coin twice and let X be the number of heads. Find and then sketch the cumulative distribution function (cdf) $F(x)$.

   Let the inverse cdf be defined as

   $$F^{-1}(t) = \min\{x : F(x) > t\} \tag{1}$$

   for any $t \in [0, 1)$.

   (b) Find $F^{-1}(0.5)$ and $F^{-1}(3/4)$.

   (c) Find $F^{-1}(0.5)$ and $F^{-1}(3/4)$ when the inverse cdf is now defined to be

   $$F^{-1}(t) = \min\{x : F(x) \geq t\}. \tag{2}$$

   Are both definitions of the inverse correct? (Hint: Find the probability mass function of $F^{-1}(U)$ where U is a uniform random variable in the interval $[0, 1)$.)

   Note: The actual definition of the inverse of the cdf is (2). The the definition in (1) has been introduced here for a learning/teaching purpose only.

4. †The set of possible of a random experiment is given by $\Omega = \{\omega_1, \omega_2, \ldots, \omega_M\}$. Show the following results using the Axioms of Probability and/or Venn diagrams.

   (a) For arbitrary events F and G, $P(F \cup G) = P(F) + P(G) - P(F \cap G)$.

   (b) Total Probability:

   $$P(G) = \sum_{i=1}^{n} P(G|F_i)P(F_i),$$

   where G is an arbitrary event and events $\{F_1, F_2, \ldots, F_n\}$ are mutually exclusive and exhaustive, i.e. $F_1 \cup F_2 \cup \ldots \cup F_n = \Omega$ and $F_i \cap F_j = 0$ for all $i \neq j$.

5. †The cdf of the random variable X is

   $$F(x) = \begin{cases} 0 & -\infty < x \leq 0 \\ 1 - e^{-x} & 0 \leq x < \infty \end{cases}$$

   Find:

   (a) $\Pr(X > 0.5)$, i.e. the probability that $X > 0.5$.

   (b) $\Pr(X \leq 0.25)$.

   (c) $\Pr(0.3 < X \leq 0.7)$.

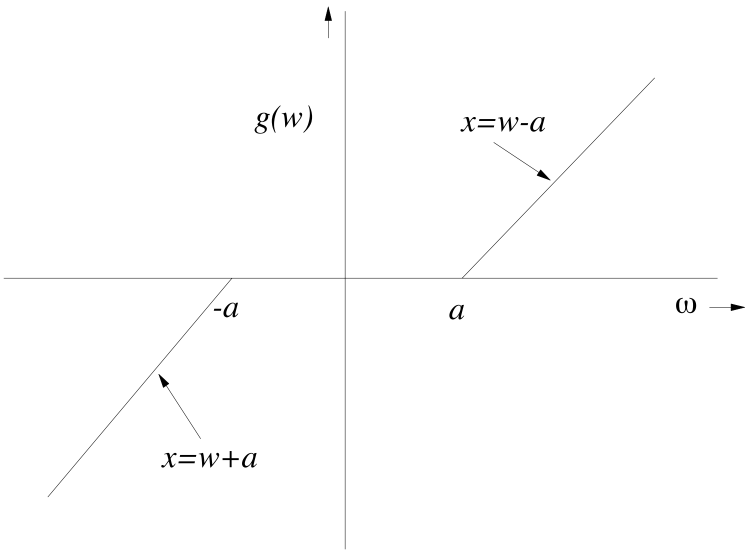

Figure 1: Function $g(\omega)$.

6. A real-valued random variable U has a Gaussian distribution with mean zero and variance σ^2. A new random variable $X = g(U)$ is defined as a function of U. Determine and sketch the probability density function of X when $g(.)$ takes the following forms:

   (a) $g(U) = U$.

   (b) $g(U) = |U|$.

   (c) $g(U) = U^2$.

   (d) $g(U)$ is as shown in the figure.

7. Let random variables X and Y have means μ_X and μ_Y and variances σ_X^2 and σ_Y^2 respectively. Define the *covariance*

$$\mathrm{Cov}(X,Y) = \mathbb{E}\left\{(X-\mu_X)(Y-\mu_Y)\right\} = \mathbb{E}\left\{XY\right\} - \mu_X\mu_Y$$

   and the *correlation* by

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y}.$$

   Let $Y = aX + b$ for constants a, b and calculate $\rho(X,Y)$.

8. An archer measures the accuracy of shots in terms of x and y coordinates relative to the centre of the bullseye. It is found from many measurements that the x- and y-values of the shots are independent and normally distributed with mean zero and standard deviation σ.

   Write down the joint probability density function for the x and y measurements and write down an integral expression which gives the probability that x lies between a and b and y lies between c and d.

   Show that the cdf for R, the radial distance of shots from the centre of the bullseye, is given by:

$$F(r) = 1 - \exp(-r^2/(2\sigma^2)), \qquad 0 \le r < \infty.$$

   Determine the pdf for R. What type of distribution is this?

9. †Let joint probability density function of random variables X and Y be $f_{X,Y}(x, y) = 1$ if $x, y \in [0, 1]$ and $f_{X,Y}(x, y) = 0$ otherwise. Show X and Y are independent.

10. Two random variables X and Y have a joint probability density function (pdf) given by

$$f_{X,Y}(x,y) = \begin{cases} kxy & 0 \leq x \leq 1, \qquad 0 \leq y \leq 1 \\ 0 & \text{Otherwise} \end{cases}$$

Determine:

(a) The value of k which makes this a valid pdf.

(b) The probability of the event $X \leq 1/2$ AND $Y > 1/2$.

(c) The marginal densities $f_X(x)$ and $f_Y(y)$.

(d) The conditional density $f_{Y|X}(y|x)$. Determine whether X and Y are independent.

11. A bivariate Gaussian pdf is

$$f_{X,Y}(x,y) = \frac{1}{2\pi\left|\Sigma\right|^{1/2}} \exp\left(-\frac{1}{2}\left[x - m_1, y - m_2\right]\Sigma^{-1}\left[x - m_1, y - m_2\right]^T\right)$$

where $\Sigma = \begin{bmatrix} \sigma_1^2 & \rho \\ \rho & \sigma_2^2 \end{bmatrix}$ is the called covariance matrix and vector (m_1, m_2) is called the mean vector.

(a) Show the conditional pdf $f_{X|Y}(x|y)$ is a Gaussian pdf with

$$\text{mean } m_1 + \frac{\rho}{\sigma_2^2}(y - m_2) \qquad \text{and variance } \sigma_1^2 - \frac{\rho^2}{\sigma_2^2}.$$

(Hint: write $f_{X,Y}(x, y) = g(x, y)h(y)$ where the function $g(x, y)$ must contain all the x and xy terms of $f_{X,Y}$ and then use the definition of the conditional pdf. Also, you may first attempt the question for $m_1 = m_2 = 0$.)

(b) Find $f_Y(y)$ by using the definition of the conditional pdf.

Answers:

Q 1 Let S = spam, NS = not spam, F = contains word free, NF = does not contain free.

$$\Omega = \{(S,F),(S,NF),(NS,F),(NS,NF)\}.$$

$$P(\{(S,F)\}) = 0.8 \times 0.25,\ P(\{(NS,F)\}) = 0.1 \times 0.75,\ P(\{(NS,NF)\}) = 0.75 \times 0.9.$$

Q 3 (a)

$$F(x) = \begin{cases} 0 & x < 0 \\ 1/4 & 0 \le x < 1 \\ 3/4 & 1 \le x < 2 \\ 1 & 2 \le x. \end{cases}$$

(b)

$$F^{-1}(0.5) = \min\{x : F(x) > 0.5\} = \min[1,\infty) = 1,$$
$$F^{-1}(3/4) = \min\{x : F(x) > 3/4\} = \min[2,\infty) = 2.$$

(c) $F^{-1}(0.5) = 1$. $F^{-1}(3/4) = 1$.

Q 5 (a) 0.607. (b) 0.221. (c) 0.244.

Q 6 (a) $\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-1}{2\sigma^2}x^2\right)$, $(-\infty < x < +\infty)$.
(b) $\frac{2}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-1}{2\sigma^2}x^2\right)$, $(0 \le x < +\infty)$.
(c) $\frac{1}{\sqrt{2\pi x\sigma^2}}\exp\left(\frac{-1}{2\sigma^2}x\right)$, $(0 \le x < +\infty)$.
(d)

$$\begin{cases} \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-1}{2\sigma^2}(x-a)^2\right), & (-\infty < x < 0) \\ (2\Phi(a/\sigma)-1)\delta(x), & (x = 0) \\ \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-1}{2\sigma^2}(x+a)^2\right), & (0 < x < \infty) \end{cases}$$

where $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}\exp(-u^2/2)du$.

Q 7 $\rho(X,Y) = a\sigma_X^2/\left(\sigma_X \times |a|\sigma_X\right) = a/|a|$.

Q 10 (a) 4. (b) 3/16. (c) $f_X(x) = \begin{cases} 2x & 0 \le x \le 1 \\ 0 & \text{Otherwise} \end{cases}$, $\quad f_Y(y) = \begin{cases} 2y & 0 \le x \le 1 \\ 0 & \text{Otherwise} \end{cases}$

(d) $\begin{cases} 2y & 0 \le y \le 1 \\ 0 & \text{Otherwise} \end{cases}$ $\quad X$ and Y independent.

S.S.Singh, September 2016