ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 4 — Advanced Digital Communications
Homework 1 — October 8, 2010

PROBLEM 1. $p_{VW}(v, w)$.

(a)

$$\begin{aligned}
E[V+W] &= \iint (v+w)\, p_{VW}(v,w)\, dv\, dw && (1)\\
&= \iint (vp_{VW}(v,w) + wp_{VW}(v,w))\, dv\, dw && (2)\\
&= \iint vp_{VW}(v,w)\, dv\, dw + \iint wp_{VW}(v,w)\, dv\, dw && (3)\\
&= \int v \int p_{VW}(v,w)\, dw\, dv + \int w \int p_{VW}(v,w)\, dv\, dw && (4)\\
&= \int vp_V(v)\, dv + \int wp_W(w)\, dw && (5)\\
&= E[V] + E[W] && (6)
\end{aligned}$$

(b)

$$\begin{aligned}
E[V \cdot W] &= \iint (v \cdot w)\, p_{VW}(v,w)\, dv\, dw && (7)\\
&= \iint (v \cdot w)\, p_V(v)\, p_W(w)\, dv\, dw && (8)\\
&= \int vp_V(v)\, dv \cdot \int wp_W(w)\, dw && (9)\\
&= E[V] \cdot E[W] && (10)
\end{aligned}$$

(c) Assume $V = W$ and $\Pr(V = 1) = \Pr(V = -1) = \frac{1}{2}$. We compute $E[V] = E[W] = 0$ and $E[VW] = 1$, so $E[VW] \neq E[V]E[W]$

Now suppose (V, W) takes values of $(1, 1), (1, -1), (-1, 1), (-1, -1), (0, 0)$ with equal probability $\frac{1}{5}$. Because $\Pr(W = 0|V = 1) = 0 \neq \frac{1}{5} = \Pr(W = 0)$, V and W are not independent. We compute $E[V] = E[W] = 0$ and $E[VW] = 0$, so $E[VW] = E[V]E[W]$

(d) Assume that V and W are independent and let σ_V^2 and σ_W^2 be the variances of V and W, respectively. Show that the variance of $V + W$ is given by $\sigma_{V+W}^2 = \sigma_V^2 + \sigma_W^2$.

$$\begin{aligned}
\sigma_{V+W}^2 &= E\left[(V+W)^2\right] - E[V+W]^2 && (11)\\
&= E[V^2] + E[W^2] + 2E[VW] - (E[V] + E[W])^2 && (12)\\
&= E[V^2] + E[W^2] + 2E[V]E[W] - E[V]^2 - E[W]^2 - 2E[V]E[W] && (13)\\
&= E[V^2] - E[V]^2 + E[W^2] - E[W]^2 && (14)\\
&= \sigma_V^2 + \sigma_W^2 && (15)
\end{aligned}$$

Problem 2.

(a)

$$\begin{aligned}
\sum_{n>0} \Pr(N \geq n) &= \sum_{n=1}^{\infty} \sum_{m=n}^{\infty} \Pr(N = m) && (16)\\
&= \sum_{m=1}^{\infty} \sum_{n=1}^{m} \Pr(N = m) && (17)\\
&= \sum_{m=1}^{\infty} m \Pr(N = m) && (18)\\
&= E[N] && (19)
\end{aligned}$$

(b)

$$\begin{aligned}
\int_0^{\infty} \Pr(x \geq a)\, da &= \int_0^{\infty} \int_a^{\infty} f_x(t)\, dt\, da && (20)\\
&= \int_0^{\infty} \int_0^{t} f_x(t)\, da\, dt && (21)\\
&= \int_0^{\infty} t\, f_x(t)\, dt && (22)\\
&= E[X] && (23)
\end{aligned}$$

(c) The main point is to note that $G(t) = P(X \geq t)$ is a non-increasing function of t. So for any fixed value of $a > 0$, the rectangle between point $(0, 0)$ and $(a, G(a))$ lies below the function $G(t)$. In conclusion, it follows from the discussion above that

$$aG(a) \leq \int_0^a G(a)\, dt \leq \int_0^a G(t)\, dt \leq \int_0^{\infty} G(t)\, dt,$$

which means

$$a \Pr(X \geq a) \leq E[X]$$

(d) Assume

$$X = (Y - E[Y])^2 \qquad X \geq 0$$

Using part (c), we have

$$a \Pr(X \geq a) \leq E[X].$$

Therefore, one could conclude that

$$a \Pr\big((Y - E[Y])^2 \geq a\big) \leq E\big((Y - E[Y])^2\big).$$

Setting $b = \sqrt{a}$, we have

$$\Pr\big(|Y - E[Y]| \geq b\big) = \Pr\big((Y - E[Y])^2 \geq b^2\big) \leq \frac{E\big((Y - E[Y])^2\big)}{b^2} = \frac{\sigma_Y^2}{b^2}.$$

Problem 3.

(a) $\Pr(X_1 \leq X_2) = \frac{1}{2}$. We know because of independence we have, $f_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1)f_{X_2}(x_2)$, and we want to find the probability of x_1 being minimum of two. This event partitions the probability space into two equal sub-sets, the other one is x_2 being the minimum of the two. The only problem is the boundary line $x_1 = x_2$, which we assume is a part of first sub-set, but because f_x is a continuous random variable the line $x_1 = x_2$ has zero probability mass and because $f_{X_1}(x_1)f_{X_2}(x_2)$ is symmetric with respect to the line $x_1 = x_2$, we conclude that the event $\min(x_1, x_2) = x_1$ partitions the whole probability space into two equally probable regions.

(b) $\Pr(X_1 \leq X_2; X_1 \leq X_3) = \frac{1}{3}$; We follow the exact same argument as the part (a), this time the probability space is partitioned into three equally probable sub-sets, in each of sub-sets one of the three random variable is minimum.

(c) Similar to last parts, we can show that

$$\Pr(X_1 \leq X_2; X_1 \leq X_3; \ldots; X_1 \leq X_{n-1}; X_1 \leq X_n) = \frac{1}{n}$$

and

$$\Pr(X_1 \leq X_2; X_1 \leq X_3; \ldots; X_1 \leq X_{n-1}) = \frac{1}{n-1}$$

We know

$$\begin{aligned}
\Pr(N = n) &= \Pr(X_1 \leq X_2; X_1 \leq X_3; \ldots; X_1 \leq X_{n-1}; X_1 > X_n) && (24)\\
&= \Pr(X_1 \leq X_2; X_1 \leq X_3; \ldots; X_1 \leq X_{n-1}) \\
&\quad - \Pr(X_1 \leq X_2; X_1 \leq X_3; \ldots; X_1 \leq X_{n-1}; X_1 \leq X_n) && (25)\\
&= \frac{1}{n-1} - \frac{1}{n} = \frac{1}{n^2 - n}. \qquad n > 1 && (26)
\end{aligned}$$

Using properties of telescopic series, we conclude

$$\begin{aligned}
\Pr(N \geq n) &= \sum_{m=n}^{\infty} \Pr(N = m) && (27)\\
&= \frac{1}{n-1} - \frac{1}{n} + \frac{1}{n} - \frac{1}{n+1} + \ldots && (28)\\
&= \frac{1}{n-1}. \qquad n \geq 2 && (29)
\end{aligned}$$

(d) We use part (a) of Problem 2.

$$E(N) = \sum_{n>0} \Pr(N \geq n) = \sum_{n>1} \frac{1}{n-1} \to \infty$$

(We know that series $\frac{1}{n}$ is divergent.)

(e) The symmetry of the $f_{X_1}(x_1)f_{X_2}(x_2)$ still holds because of independence but in the discrete case it is possible to put some probability mass on the line $x_1 = x_2$. Therefore in the discrete case the event $x_1 \leq x_2$ does not partition the whole probability space into two equally probable sub-spaces. The same as before we can conclude that $\Pr(X_1 < X_2) = \Pr(X_2 < X_1)$. We know $\Pr(X_1 < X_2) + \Pr(X_1 = X_2) + \Pr(X_2 < X_1) = 1$. From these two we conclude that $\Pr(X_1 \leq X_2) \geq \frac{1}{2}$. Similarly we conclude that

$$\Pr(X_1 \leq X_2; X_1 \leq X_3; \ldots; X_1 \leq X_{n-1}; X_1 \leq X_n) \geq \frac{1}{n}.$$

Following the steps in part (d), we can show that

$$E(N) \geq \sum_{n>1} \frac{1}{n-1} \to \infty$$

PROBLEM 4. Let's consider the case where $n = 2$ first, we have

$$P(Z = 0) = P(X_1 \oplus X_2 = 0) = P(X_1 = 0, X_2 = 0) + P(X_1 = 1, X_2 = 1) = \frac{1}{2},$$

in which we used independence of X_1 and X_2.

By induction, one could easily show that for arbitrary n, we have

$$P(Z = 0) = \frac{1}{2}.$$

(a)

$$\begin{aligned} P(Z = z|X_1 = x_1) &= P(X_1 \oplus X_2 \oplus \cdots \oplus X_n = z|X_1 = x_1) & (30) \\ &= P(X_2 \oplus \cdots \oplus X_n = z \oplus x_1|X_1 = x_1) & (31) \\ &= P(X_2 \oplus \cdots \oplus X_n = z \oplus x_1) & (32) \\ &= \frac{1}{2} = P(Z = z) & (33) \end{aligned}$$

in (32) we used that X_i's are independent. We conclude that Z is independent of X_1

(b)

$$\begin{aligned} P(Z = z|X_1, \ldots, X_{n-1} = x_1, \ldots, x_{n-1}) &= & (34) \\ P(X_1 \oplus X_2 \oplus \cdots \oplus X_n = z|X_1, \ldots, X_{n-1} = x_1, \ldots, x_{n-1}) &= & (35) \\ P(X_n = z \oplus x_1 \oplus \cdots \oplus x_{n-1}|X_1, \ldots, X_{n-1} = x_1, \ldots, x_{n-1}) &= & (36) \\ P(X_n = z \oplus x_1 \oplus \cdots \oplus x_{n-1}) &= & (37) \\ &= \frac{1}{2} & (38) \\ &= P(Z = z) & (39) \end{aligned}$$

in (37) we used that X_i's are independent. We conclude that Z is independent of $X_1, \ldots, X_{n-1}$.

(c) No, Z is a deterministic function of $X_1, \ldots, X_n$, which means

$$P(Z = z|X_1, \ldots, X_n = x_1, \ldots, x_n)$$

is either 0 or 1 depending on the values of $x_1, \ldots, x_n$ and z.

(d) Suppose $\Pr(X_i = 1) = \frac{3}{4}$, we have

$$P(Z = 0) = P(X_1 \oplus X_2 = 0) = P(X_1 = 0, X_2 = 0) + P(X_1 = 1, X_2 = 1) = \frac{9+1}{16} = \frac{5}{8}.$$

but

$$\begin{aligned} P(Z = 0|X_1 = 0) &= P(X_1 \oplus X_2 = 0|X_1 = 0) & (40) \\ &= P(X_2 = 0|X_1 = 0) & (41) \\ &= \frac{1}{4} \neq \frac{5}{8} = P(Z = 0), & (42) \end{aligned}$$

in which we used that X_1 and X_2 are independent. We conclude that Z is not independent of X_1.

PROBLEM 5. (1) Let D_0, D_1 be the MAP decision regions for hypotheses 0 and 1 when the a-priori probabilities are $(\pi_0, 1-\pi_0)$. Similarly, let D'_0, D'_1 be the MAP decision regions for hypotheses 0 and 1 when the a-priori probabilities are $(\pi'_0, 1-\pi'_0)$, and D''_0, D''_1 be the MAP decision regions for hypotheses 0 and 1 when the a-priori probabilities are $(\pi''_0, 1-\pi''_0)$, where $\pi''_0 = \lambda\pi_0 + (1-\lambda)\pi'_0$. Thus

$$\begin{aligned} V(\pi_0) &= \pi_0 p_0(D_1) + (1-\pi_0)p_1(D_0), \\ V(\pi'_0) &= \pi_0 p_0(D'_1) + (1-\pi_0)p_1(D'_0), \\ V(\pi''_0) &= \pi_0 p_0(D''_1) + (1-\pi_0)p_1(D''_0), \end{aligned}$$

(2) Since the MAP rule minimizes the error probability, using any other decision regions in any of the above will increase the probability of error. So,

$$\begin{aligned} V(\pi_0) &\le \pi_0 p_0(D''_1) + (1-\pi_0)p_1(D''_0), \\ V(\pi'_0) &\le \pi'_0 p_0(D''_1) + (1-\pi'_0)p_1(D''_0). \end{aligned}$$

Multiplying the first by λ and the second by $(1-\lambda)$ and adding we get the desired result:

$$\begin{aligned} \lambda V(\pi_0) + (1-\lambda)V(\pi'_0) &\le (\lambda\pi_0 + (1-\lambda)\pi'_0)p_0(D''_1) \\ &\quad +(1-(\lambda\pi_0 + (1-\lambda)\pi'_0))p_1(D''_0) \\ &= V(\lambda\pi_0 + (1-\lambda)\pi'_0) \end{aligned} \quad (43)$$

PROBLEM 6. We define

$$C(x_i) = 2\sigma^2 \log \Pr(x_i)$$

It is easy to show that for the optimal decision maker (MAP) in Gaussian noise, the detector finds x_i so that

$$\langle x_i, x_i\rangle - 2\langle y, x_i\rangle - C(x_i)$$

is minimized.

We know the following for any $j \ne i$

$$\langle x_i, x_i\rangle - 2\langle y_1, x_i\rangle - C(x_i) \le \langle x_j, x_j\rangle - 2\langle y_1, x_j\rangle - C(x_j) \quad (44)$$

$$\langle x_i, x_i\rangle - 2\langle y_2, x_i\rangle - C(x_i) \le \langle x_j, x_j\rangle - 2\langle y_2, x_j\rangle - C(x_j). \quad (45)$$

Now let us consider the following,

$$\begin{aligned} \langle x_i, x_i\rangle - 2\langle \alpha y_1 + (1-\alpha)y_2, x_i\rangle - C(x_i) &= \langle x_i, x_i\rangle - 2\alpha\langle y_1, x_i\rangle \\ &\quad -2(1-\alpha)\langle y_2, x_i\rangle - C(x_i) \\ &= \alpha\big[\langle x_i, x_i\rangle - 2\langle y_1, x_i\rangle - C(x_i)\big] + \\ &\quad (1-\alpha)\big[\langle x_i, x_i\rangle - 2\langle y_2, x_i\rangle - C(x_i)\big] \\ &\le \alpha\big[\langle x_j, x_j\rangle - 2\langle y_1, x_j\rangle - C(x_j)\big] + \\ &\quad (1-\alpha)\big[\langle x_j, x_j\rangle - 2\langle y_2, x_j\rangle - C(x_j)\big]. \end{aligned}$$

In the last step we used 44 and 45. We conclude

$$\langle x_i, x_i\rangle - 2\langle \alpha y_1 + (1-\alpha)y_2, x_i\rangle - C(x_i) \le \langle x_j, x_j\rangle - 2\langle \alpha y_1 + (1-\alpha)y_2, x_j\rangle - C(x_j)$$

for all $j \ne i$. Therefore, the decoder decodes $\alpha y_1 + (1-\alpha)y_2$ as x_i.