①

(a) $\leq$.

By law of total probability,

$P(A_1)\cdot P(B|A_1) + P(A_2)\cdot P(B|A_2) = P(A_1,B) + P(A_2,B) = *$

since $P(A_1) \cup P(A_2) = \Omega$, $* \geq P(B)$.

(b) $\leq$

Entropy of a Bernoulli R.V. ≤ 1.

(c) $=$

$P(A) = E[\mathbb{1}_{\{x \in A\}}] = E[g(x)]$

(d) $=$

$E[X^a Y^b] = \sum_x \sum_y x^a y^b p(x,y)$, def of $E[\cdot]$

$\qquad = \sum_x \sum_y x^a y^b p(x) p(y)$, $x \perp y$

$\qquad = \sum_x x^a p(x) \sum_y y^b p(y)$

$\qquad = E[x] E[Y]$

(e) $\leq$.

By Markov inequality,

$P((X+Y)^2 \geq 16) \leq \dfrac{E[(x+y)^2]}{16} = \dfrac{4+4+0}{16} = \dfrac{1}{2}$

(f) $=$

$Cov(X_i, X_j) = E[X_i X_j] - E[X_i] E[X_j]$

②

(a) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\, dx\, dy = 1$

$\int_0^1 \int_0^x cx \, dy\, dx = c \int_0^1 x(x-0)\, dx$

$\qquad = c\int_0^1 x^2 dx = c [\frac{1}{3}x^3]_0^1 = \frac{1}{3}c = 1$ $\therefore c = 3$

(b) $f(y) = \int_y^0 3x\, dx = (\frac{3}{2}x^2]_y^{\blacksquare} = \frac{3}{2}(1-y^2)$ for $0 \leq y \leq 1$

$f(x|y) = \dfrac{f(x,y)}{f(y)} = \dfrac{3x}{\frac{3}{2}(1-y^2)} = \dfrac{2x}{1-y^2}$

$\therefore E[X|Y=y] = \int_{-\infty}^{\infty} x f(x|y)\, dx = \int_y^1 \dfrac{2x^2}{1-y^2} = \frac{2}{3}\frac{1}{1-y^2}[x^3]_y^1$

$\qquad = \dfrac{2(1-y^3)}{3(1-y^2)}$

(c) $f(x) = \int_0^x 3x\, dy = 3x^2$

$f(x)\cdot f(y) = 3x^2 \cdot \frac{3}{2}(1-y^2) = \frac{9}{2}x^2(1-y^2) \neq f(x,y)$

$\therefore x \not\perp y$

③

(a) No. For example, for $p(x_1|y=0)$, the value of x_1 depends on x_2.

(b) $\hat{y} = \underset{y}{\arg\max}\; p(y|x)$

	$x_2=0$	$x_2=1$
$x_1=0$	0 or 1	0
$x_2=1$	1	0
$x \neq 2$	1	1

(c) true risk $= \underset{n\to\infty}{\lim} \frac{1}{n} \sum_i \ell(\hat{y}_i, y_i) = E[\ell(\hat{y}_i, y_i)]$

$\qquad = E[\mathbb{1}_{\{\hat{y}_i = y_i\}}]$

(d) There're only 2 cases where classification is not deterministic:

$y=0, x_1 = 0.5, x_2 = 0$ $\Rightarrow P = 0.1$

$y=1, x_1=1, x_2=1$ $\Rightarrow P(y=1, x_1=1, x_2=1) = 0.1$

True risk $= .2$.