2.1 X = one child, Y = other child

X	Y	$P(X,Y)$
G	G	1/4
G	B	1/4
B	G	1/4
B	B	1/4

(a) (2 Point) Let N_g = number of girls, N_b = number of boys with constraint $N_g + N_b = 2$

$$P(N_g=1 \mid N_b \geq 1) = \frac{P(N_b \geq 1 \mid N_g = 1)\, P(N_g = 1)}{P(N_b \geq 1)} = \frac{1 \times 1/2}{3/4} = 2/3$$

(b) (2 point) Let Y = the identity of the observed child

X = identity of the other child

Then $$P(X=g \mid Y=b) = \frac{P(Y=b \mid X=g)\, P(X=g)}{P(Y=b)} = \frac{1/2 \times 1/2}{1/2} = 1/2$$

2.3 (2 points) Variance of a sum

$$\mathrm{Var}[X+Y] = E[(X+Y)^2] - (E[X+Y])^2 \quad (\because \mathrm{Var}(Z) = E[Z^2] - (E[Z])^2)$$

$$= E[X^2+Y^2+2XY] - (E[X]+E[Y])^2 \quad (\because \text{Expectation is linear})$$

$$= E[X^2] + E[Y^2] + 2E[XY] - (E[X])^2 - (E[Y])^2 - 2E[X]E[Y]$$

$$= \underline{E[X^2] - (E[X])^2} + \underline{E[Y^2] - (E[Y])^2} + \underline{2E[XY] - 2E[X]E[Y]}$$

$$= \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\,\mathrm{Cov}[X,Y]$$

2.6

(a) (3 points) Baye's rule gives

$$P(H \mid E_1=e_i, E_2=e_j) = \frac{P(E_1=e_i, E_2=e_j \mid H)\, P(H)}{P(E_1, E_2)}$$

Hence (ii) is sufficient (we even don't need $P(e_1, e_2)$)

(i), (iii) are insufficient

(b) (3 Points) If $E_1 \perp E_2 \mid H$ (E_1 and E_2 are conditionally independent given H)

then $$P(H \mid E_1=e_i, E_2=e_j) = \frac{P(E_1=e_i \mid H)\, P(E_2=e_j \mid H)\, P(H)}{P(E_1=e_i, E_2=e_j)}$$

(i) and (ii) are obviously sufficient

(iii) is also sufficient, because we can compute $P(E_1, E_2)$ for normalization

2.12 Expressing mutual information in terms of entropies

(2 points)

$$I[X,Y] = \sum_{x,y} P(x,y) \log \frac{P(x,y)}{P(x)P(y)}$$

$$= \sum_{x,y} P(x,y) \log \frac{P(x|y)P(y)}{P(x)P(y)}$$

$$= -\sum_{x,y} P(x,y) \log P(x) + \sum_{x,y} P(x,y) \lg (x|y)$$

marginalization

$$= -\sum_{x} P(x) \log P(x) - \left(-\sum_{x,y} P(x,y) \log (x|y)\right)$$

$$= H[X] - \left(-\sum_{y} P(y) \sum_{x} P(y|x) \log (x|y)\right)$$

$$= H[x] - H[x|y]$$

and $I[X,Y] = H(Y) - H[Y|X]$ by symmetry

2.16 $\quad \text{Beta}(x|a,b) = \frac{1}{B[a,b]} x^{a-1}(1-x)^{b-1}$

(2 points)

mode $= x$ where $\text{Beta}(x|a,b)$ has maximum value

Hence using simple calculus we have

$$\frac{d\,\text{Beta}(x|a,b)}{dx} = \frac{1}{B[a,b]}\left[-x^{(a-1)}(b-1)(1-x)^{b-2} + (a-1)x^{a-2}(1-x)^{b-1}\right] = 0$$

$$\Rightarrow \frac{x^{a-2}(1-x)^{b-2}}{B[a,b]}\left[-(b-1)x + (a-1)(1-x)\right] = 0$$

$$\Rightarrow \left[(a-1) - (b-1+a-1)x\right) = 0$$

$$\Rightarrow x = \frac{a-1}{a+b-2}$$

mean (2 points) $\quad E[\cancel{x}] = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 x\, x^{a-1}(1-x)^{b-1}\, dx = \frac{\Gamma(a+b)\Gamma(a+1)\Gamma(b)}{\Gamma(a)\Gamma(b)\Gamma(a+1+b)} = \frac{a}{a+b}$

For variance first observe that $E[x^2] = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 x^2 \left(x^{a-1}(1-x)^{b-1}\right) dx$

$= \cancel{\frac{\Gamma(a+b)\,\Gamma(a+2+b)}{\Gamma(a)\,\Gamma(b)\,\Gamma(a+b)\,\Gamma(b)}}$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+2)\Gamma(b)}{\Gamma(a+2+b)} =$$

$$= \frac{(a+1)\,a}{(a+1+b)\,(a+b)}$$

$\left(\because \Gamma(z+1) = z\,\Gamma(z) \right.$
<u>Note</u> gamma function behave like factorial with its argument shifted down by 1 $)$

Hence $Var[x] = \mathbb{E}[\theta^2] - (\mathbb{E}[\theta])^2 = \frac{(a+1)a}{(a+1+b)(a+b)} - \left(\frac{a}{a+b}\right)^2$
(2 points)

$$= \frac{a\,b}{(a+b)^2\,(a+b+1)}$$

$$\boxed{\begin{aligned} \text{Total Points} &= (2+2) + (2) + (3+3) +2 + (2+2+2) \\ &= 4 + 2 + 6 +2 + 6 \\ &= 18 +2 = 20 \end{aligned}}$$

1   3   6   12   16