1) (a) $f_{(X,Y)}$ is obtained by dividing each entry in the table by the total 205 obtaining

$X \backslash Y$	1	2	3	4
1	33/205	26/205	10/205	20/205
2	21/205	10/205	13/205	12/205
3	16/205	18/205	16/205	20/205

(b) The conditional distributions are obtained by taking the original table and dividing each row by the row total.

$f_Y(\cdot\mid X=1)$	0.371	0.292	0.112	0.225
$f_Y(\cdot\mid X=2)$	0.375	0.179	0.232	0.214
$f_Y(\cdot\mid X=3)$	0.229	0.257	0.229	0.286

(c) There is a relationship between X and Y because the conditional distributions are different.
Plotting the conditional distributions reveals no discernible pattern to this relationship.

2 $\Rightarrow$] We have that

$$f_{Y|X}(y|x) = \frac{\#\{\omega : X(\omega)=x,\ Y(\omega)=y\}}{\#\{\omega : X(\omega)=x\}}$$

$$= \frac{\#\{\omega : X(\omega)=x,\ Y(\omega)=y\}}{\#(\Omega)} \Big/ \frac{\#\{\omega : X(\omega)=x\}}{\#(\Omega)}$$

$$= f_{(X,Y)}(x,y) / f_X(x)$$

$$\overset{\text{hypothesis}}{=} \frac{f_X(x) f_Y(y)}{f_X(x)} = f_Y(y)$$

and therefore $f_{Y|X}(\cdot|x)$ doesn't change as we change x and there is no relationship.

$\Leftarrow$] Let $g(y) = f_{Y|X}(y|x)$ as this doesn't depend on x by hypothesis. Since (as argued above)

$$f_{(X,Y)}(x,y) = f_{Y|X}(y|x) f_X(x)$$

$$= g(y) f_X(x)$$

and so $f_Y(y) = \sum_x f_{(X,Y)}(x,y)$

$$= \sum_x g(y) f_X(x) = g(y)\left(\sum_x f_X(x)\right)$$

$= g(y)$ since $\sum_x f_X(x) = \sum_x \frac{\#\{\omega : X(\omega)=x\}}{\#(\Omega)} = 1$

(3) $\sum_{x\in(c_i,c_j]} f_X(x) = \sum_{x\in(c_i,c_{i+1}]} f_X(x) + \sum_{x\in(c_{i+1},c_{i+2}]} f_X(x) + \cdots + \sum_{(c_{j-1},c_j]} f_X(x)$

and $\sum_{x\in(c_i,c_{i+1}]} f_X(x) = \frac{\#\{\omega : X(\omega)\in(c_i,c_{i+1}]\}}{\#(\Omega)}$. Now

$$\int_{c_i}^{c_j} f_X^*(x)\,dx = \int_{c_i}^{c_{i+1}} f_X^*(x)\,dx + \int_{c_{i+1}}^{c_{i+2}} f_X^*(x)\,dx + \cdots + \int_{c_{j-1}}^{c_j} f_X^*(x)\,dx$$

and $\int_{c_i}^{c_{i+1}} f_X^*(x)\,dx = \int_{c_i}^{c_{i+1}} \frac{\#\{\omega : X(\omega)\in(c_i,c_{i+1}]\}}{\#(\Omega)(c_{i+1}-c_i)}\,dx$

$$= \frac{\#\{\omega : X(\omega)\in(c_i,c_{i+1}]\}}{\#(\Omega)(c_{i+1}-c_i]}\int_{c_i}^{c_{i+1}} dx$$

$$= \frac{\#\{\omega : X(\omega)\in(c_i,c_{i+1}]\}}{\#(\Omega)}\frac{\cancel{(c_{i+1}-c_i)}}{\cancel{(c_{i+1}-c_i)}} = \sum_{x\in(c_i,c_{i+1}]} f_X(x)$$

which establishes the result.

(4) (a) We have that

$$f_{Y|X}(y|x) = f_{(X,Y)}(x,y)\,/\,f_X(x)$$

and using $X\sim N(\mu_1,\sigma_1^2)$ this implies

$f_{Y|X}(y|x)$ using $\begin{pmatrix}\sigma_1^2 & \sigma_1\sigma_2\rho\\ \sigma_1\sigma_2\rho & \sigma_2^2\end{pmatrix}^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)}\begin{pmatrix}\sigma_2^2 & -\sigma_1\sigma_2\rho\\ -\sigma_1\sigma_2\rho & \sigma_1^2\end{pmatrix}$

$$= \frac{(2\pi)^{-1}\left(\sigma_1^2\sigma_2^2(1-\rho^2)\right)^{-\frac{1}{2}}\exp\left\{\frac{-1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\right\}}{(2\pi)^{-\frac{1}{2}}(\sigma_1^2)^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2\right\}}$$

$$= (2\pi)^{-\frac{1}{2}} \left(\sigma_2^2(1-\rho^2)\right)^{-\frac{1}{2}} \times$$

$$\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2 - (1-\rho^2)\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right]\right\}$$

and $[\ ] = \rho^2\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2$

$$= \left[\left(\frac{y-\mu_2}{\sigma_2}\right) - \rho\left(\frac{x-\mu_1}{\sigma_1}\right)\right]^2 = \frac{1}{\sigma_2^2}\left[y - \mu_2 - \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right]^2$$

$\therefore\ Y \mid X = x \sim N\left(\mu_2 + \frac{\rho\sigma_2}{\sigma_1}(x-\mu_1),\ \sigma_2^2(1-\rho^2)\right)$

(b) When $\rho > 0$ then X and Y are correlated so they vary together. When we condition on X, namely fix X, then we are removing the joint variation from Y leaving only the variation in Y that has nothing to do with X.