6-4 Joint Moments

Given two random variables $\mathbf{x}$ and $\mathbf{y}$ and a function $g(x, y)$ ($g\colon \mathbb{R}^2 \to \mathbb{R}$), we form the random variable $\mathbf{z} = g(\mathbf{x}, \mathbf{y})$.

- The expected value of $\mathbf{z}$ is

$$E\{\mathbf{z}\} = \begin{cases} \displaystyle\int_{-\infty}^{\infty} z f_{\mathbf{z}}(z)\, dz & \text{continuous RV} \\ \displaystyle\sum_{\ell} z_\ell P(\mathbf{z} = z_\ell) & \text{discrete RV} \end{cases}$$

- From the Law of the Unconscious Statistician

$$E\{\mathbf{z}\} = \begin{cases} \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y) f_{\mathbf{xy}}(x, y)\, dx\, dy & \text{jointly continuous RVs} \\ \displaystyle\sum_{i}\sum_{k} g(x_i, y_k) P(\mathbf{x} = x_i, \mathbf{y} = y_k) & \text{jointly discrete RVs} \end{cases}$$

- Linearity

$$E\left\{\sum_{k=1}^{n} a_k g_k(\mathbf{x}, \mathbf{y})\right\} = \sum_{k=1}^{n} a_k E\left\{g_k(\mathbf{x}, \mathbf{y})\right\}$$

  Consequences

$$E\{\mathbf{x} + \mathbf{y}\} = E\{\mathbf{x}\} + E\{\mathbf{y}\}$$

$$E\{\mathbf{xy}\} \neq E\{\mathbf{x}\}E\{\mathbf{y}\} \quad \text{in general}$$

Definitions

$\mathbf{x}$ and $\mathbf{y}$ are two random variables with

$$E\{\mathbf{x}\} = \mu_{\mathbf{x}} \qquad E\{(\mathbf{x}-\mu_{\mathbf{x}})^2\} = \sigma_{\mathbf{x}}^2$$
$$E\{\mathbf{y}\} = \mu_{\mathbf{y}} \qquad E\{(\mathbf{y}-\mu_{\mathbf{y}})^2\} = \sigma_{\mathbf{y}}^2$$

- The *covariance* $C_{\mathbf{xy}}$ is

$$C_{\mathbf{xy}} = E\{(\mathbf{x}-\mu_{\mathbf{x}})(\mathbf{y}-\mu_{\mathbf{y}})\}$$

- The *correlation coefficient* is

$$\rho_{\mathbf{xy}} = \frac{C_{\mathbf{xy}}}{\sigma_{\mathbf{x}}\sigma_{\mathbf{y}}} \qquad -1 \le \rho_{\mathbf{xy}} \le 1$$

- The *correlation* is

$$R_{\mathbf{xy}} = E\{\mathbf{xy}\}$$

- $\mathbf{x}$ and $\mathbf{y}$ are *uncorrelated* means

$$C_{\mathbf{xy}} = 0 \qquad \rho_{\mathbf{xy}} = 0 \qquad E\{\mathbf{xy}\} = E\{\mathbf{x}\}E\{\mathbf{y}\}$$

- $\mathbf{x}$ and $\mathbf{y}$ are *orthogonal* means

$$R_{\mathbf{xy}} = 0$$

- Theorem 6-5: If two random variables $\mathbf{x}$ and $\mathbf{y}$ are independent $[f_{\mathbf{xy}}(x,y) = f_{\mathbf{x}}(x)f_{\mathbf{y}}(y)]$ then they are uncorrelated.

- Variance of the sum $\mathbf{z} = \mathbf{x} + \mathbf{y}$:

$$\sigma_{\mathbf{z}}^2 = \sigma_{\mathbf{x}}^2 + 2\rho_{\mathbf{xy}}\sigma_{\mathbf{x}}\sigma_{\mathbf{y}} + \sigma_{\mathbf{y}}^2$$

- *Joint moments*

$$m_{kr} = E\{\mathbf{x}^k\mathbf{y}^r\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^k y^r f_{\mathbf{xy}}(x,y)\,dx\,dy$$

- *Joint central moments*:

$$\mu_{kr} = E\left\{(\mathbf{x}-m_{10})^k\,(\mathbf{y}-m_{01})^r\right\}$$

Comments

$\mathbf{x}$ and $\mathbf{y}$ are two random variables with

$$E\{\mathbf{x}\} = \mu_{\mathbf{x}} \qquad E\{(\mathbf{x}-\mu_{\mathbf{x}})^2\} = \sigma_{\mathbf{x}}^2$$
$$E\{\mathbf{y}\} = \mu_{\mathbf{y}} \qquad E\{(\mathbf{y}-\mu_{\mathbf{y}})^2\} = \sigma_{\mathbf{y}}^2$$

- Comment on covariance

$$C_{\mathbf{xy}} = R_{\mathbf{xy}} - \mu_{\mathbf{x}}\mu_{\mathbf{y}}$$

- Comment on jointly normal random variables $\mathbf{x}$ and $\mathbf{y}$.

  MDR's preferred form is

$$f_{\mathbf{xy}}(x,y) = \frac{1}{2\pi\sqrt{\det(C)}} \exp\left\{-\frac{1}{2}\left[(x-\mu_{\mathbf{x}}) \quad (y-\mu_{\mathbf{y}})\right] C^{-1} \begin{bmatrix} x-\mu_{\mathbf{x}} \\ y-\mu_{\mathbf{y}} \end{bmatrix}\right\}$$

  where
$$C = \begin{bmatrix} \sigma_{\mathbf{x}}^2 & r\sigma_{\mathbf{x}}\sigma_{\mathbf{y}} \\ r\sigma_{\mathbf{x}}\sigma_{\mathbf{y}} & \sigma_{\mathbf{y}}^2 \end{bmatrix}$$
  The parameter r here is defined by

$$r\sigma_{\mathbf{x}}\sigma_{\mathbf{y}} = E\left\{(\mathbf{x}-\mu_{\mathbf{x}})(\mathbf{y}-\mu_{\mathbf{y}})\right\} = C_{\mathbf{xy}}.$$

  That is, $r = \rho_{\mathbf{xy}}$.

- Comment on Theorem 6-5

  If two random variables are uncorrelated they are not necessarily independent. However, for normal random variables uncorrelatedness is equivalent to independence.

- Comment on moments

  For the determination of the joint statistics of $\mathbf{x}$ and $\mathbf{y}$ knowledge of their joint density is required. However, in many applications, only the first- and second-moments are used. These moments are determined in terms of the five parameters

$$\mu_{\mathbf{x}} \quad \mu_{\mathbf{y}} \quad \sigma_{\mathbf{x}}^2 \quad \sigma_{\mathbf{y}}^2 \quad \rho_{\mathbf{xy}}$$

  If $\mathbf{x}$ and $\mathbf{y}$ are jointly normal, then these parameters determine uniquely $f_{\mathbf{xy}}(x,y)$.

6-5 Joint Characteristic Functions

Definition

The *joint characteristic function* of the random variables $\mathbf{x}$ and $\mathbf{y}$ is

$$\Phi_{\mathbf{xy}}(\omega_1, \omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\mathbf{xy}}(x, y) e^{j(\omega_1 x + \omega_2 y)} dx\, dy$$

Properties

1. The *inversion formula* is

$$f_{\mathbf{xy}}(x, y) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi_{\mathbf{xy}}(\omega_1, \omega_2) e^{-j(\omega_1 x + \omega_2 y)} d\omega_1\, d\omega_2$$

2. The *marginal characteristic functions* are

$$\Phi_{\mathbf{x}}(\omega) = \Phi_{\mathbf{xy}}(\omega, 0) \qquad \Phi_{\mathbf{y}}(\omega) = \Phi_{\mathbf{xy}}(0, \omega)$$

3. If $\mathbf{z} = a\mathbf{x} + b\mathbf{y}$ then

$$\Phi_{\mathbf{z}}(\omega) = E\left\{e^{j(a\mathbf{x}+b\mathbf{y})}\right\} = \Phi_{\mathbf{xy}}(a\omega, b\omega).$$

4. Independence. If the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent

$$E\left\{e^{j(\omega_1 \mathbf{x} + \omega_2 \mathbf{y})}\right\} = E\left\{e^{j\omega_1 \mathbf{x}}\right\} E\left\{e^{j\omega_2 \mathbf{y}}\right\}$$

$$\Rightarrow \quad \Phi_{\mathbf{xy}}(\omega_1, \omega_2) = \Phi_{\mathbf{x}}(\omega_1)\Phi_{\mathbf{y}}(\omega_2)$$

5. Convolution. If the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent and $\mathbf{z} = \mathbf{x} + \mathbf{y}$, then

$$E\left\{e^{j\omega \mathbf{z}}\right\} = E\left\{e^{j\omega(\mathbf{x}+\mathbf{y})}\right\} = E\left\{e^{j\omega \mathbf{x}}\right\} E\left\{e^{j\omega \mathbf{y}}\right\}$$

$$\Rightarrow \Phi_{\mathbf{z}}(\omega) = \Phi_{\mathbf{x}}(\omega)\Phi_{\mathbf{y}}(\omega)$$

Convolution of pdfs $\Leftrightarrow$ multiplication of characteristic functions.

A Random Result

If $\mathbf{x}$ and $\mathbf{y}$ are jointly normal with zero mean, then

$$E\left\{\mathbf{x}^2\mathbf{y}^2\right\} = E\left\{\mathbf{x}^2\right\} E\left\{\mathbf{y}^2\right\} + 2E^2\left\{\mathbf{xy}\right\}$$

6-6 Conditional Distributions

Definition

The *conditional distribution function* of the random variable and $\mathbf{y}$ given the event M is

$$F_{\mathbf{y}|M}(y|M) = P(\{\zeta \in \mathcal{S}\colon \mathbf{y}(\zeta) \le y\}|\{\zeta \in \mathcal{S}\colon M \text{ occurs}\})$$

$$= \frac{P(\{\zeta \in \mathcal{S}\colon \mathbf{y}(\zeta) \le y\} \cap \{\zeta \in \mathcal{S}\colon M \text{ occurs}\})}{P(\{\zeta \in \mathcal{S}\colon M \text{ occurs}\})}$$

or, using the short-hand notation,

$$= \frac{P(\mathbf{y} \le y, M)}{P(M)}$$

Properties

1. $M = \mathbf{x}(\zeta) \le x$:

$$F_{\mathbf{y}|\mathbf{x}\le x}(y|\mathbf{x} = x) = \frac{F_{\mathbf{xy}}(x, y)}{F_{\mathbf{x}}(x)}$$

$$f_{\mathbf{y}|\mathbf{x}\le x}(y|\mathbf{x} \le x) = \frac{1}{F_{\mathbf{x}}(x)} \int_{-\infty}^{x} f_{\mathbf{xy}}(u, y)\, du$$

2. $M = x_1 \le \mathbf{x}(\zeta) \le x_2$:

$$F_{\mathbf{y}|x_1<\mathbf{x}\le x_2}(y|x_1 < \mathbf{x} \le x_2) = \frac{F_{\mathbf{xy}}(x_2, y) - F_{\mathbf{xy}}(x_1, y)}{F_{\mathbf{x}}(x_2) - F_{\mathbf{x}}(x_1)}$$

$$f_{\mathbf{y}|x_1<\mathbf{x}\le x_2}(y|x_1 < \mathbf{x} \le x_2) = \frac{1}{F_{\mathbf{x}}(x_2) - F_{\mathbf{x}}(x_1)} \int_{x_1}^{x_2} f_{\mathbf{xy}}(u, y)\, du$$

3. Same as 2 with $x_1 = x$ and $x_2 = x + \Delta x$. As $\Delta x \to 0$:

$$f_{\mathbf{y}|\mathbf{x}=x}(y|\mathbf{x} = x) = \frac{f_{\mathbf{xy}}(x, y)}{f_{\mathbf{x}}(x)}$$

4. The other way round

$$f_{\mathbf{x}|\mathbf{y}=y}(\mathbf{x}|\mathbf{y} = y) = \frac{f_{\mathbf{xy}}(x, y)}{f_{\mathbf{y}}(y)}$$

Bayes' Theorem and Total Probability

- Bayes's Rule

$$f_{\mathbf{x}|\mathbf{y}}(x|y) = \frac{f_{\mathbf{y}|\mathbf{x}}(y|x)f_{\mathbf{x}}(x)}{f_{\mathbf{y}}(y)} \qquad f_{\mathbf{y}|\mathbf{x}}(y|x) = \frac{f_{\mathbf{x}|\mathbf{y}}(x|y)f_{\mathbf{y}}(y)}{f_{\mathbf{x}}(x)}$$

- Total Probability Theorem

$$f_{\mathbf{y}}(y) = \int_{-\infty}^{\infty} f_{\mathbf{y}|\mathbf{x}}(y|x)f_{\mathbf{x}}(x)\,dx \qquad f_{\mathbf{x}}(x) = \int_{-\infty}^{\infty} f_{\mathbf{x}|\mathbf{y}}(x|y)f_{\mathbf{y}}(y)\,dy$$

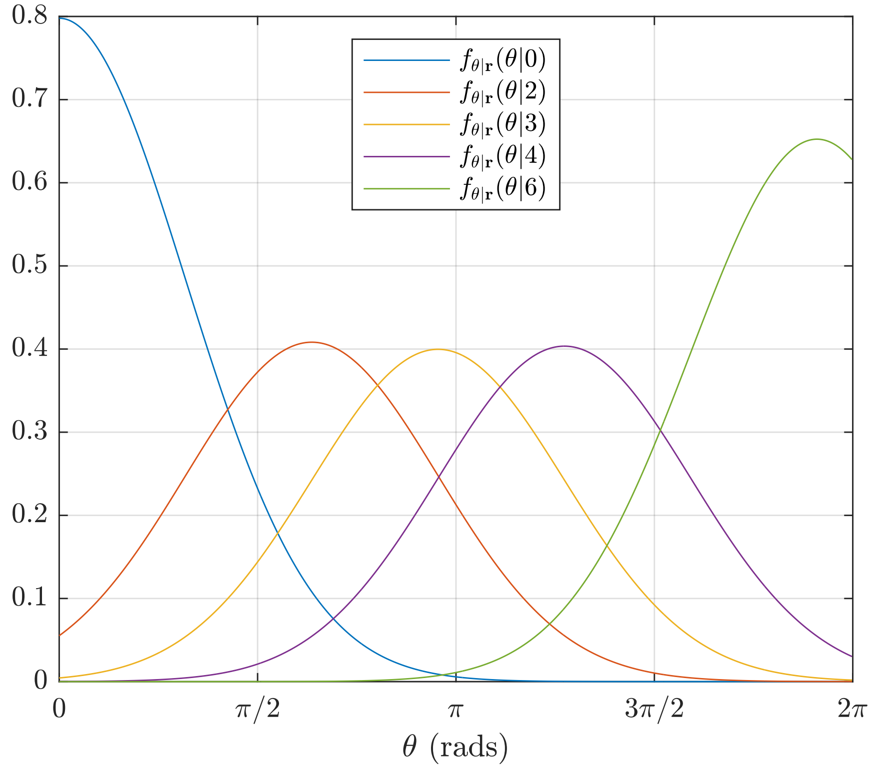
- Bayes' Theorem

$$f_{\mathbf{x}|\mathbf{y}}(x|y) = \frac{f_{\mathbf{y}|\mathbf{x}}(y|x)f_{\mathbf{x}}(x)}{\displaystyle\int_{-\infty}^{\infty} f_{\mathbf{y}|\mathbf{x}}(y|x)f_{\mathbf{x}}(x)\,dx} \qquad f_{\mathbf{y}|\mathbf{x}}(y|x) = \frac{f_{\mathbf{x}|\mathbf{y}}(x|y)f_{\mathbf{y}}(y)}{\int_{-\infty}^{\infty} f_{\mathbf{x}|\mathbf{y}}(x|y)f_{\mathbf{y}}(y)\,dy}$$

	Event A with Event B	Event A with RV $\mathbf{x}$	RV $\mathbf{x}$ with RV $\mathbf{y}$
Bayes Rule	$P(A\|B) = \frac{P(B\|A)P(A)}{P(B)}$	$P(A\|\mathbf{x}=x) = \frac{f_{\mathbf{x}\|A}(x\|A)P(A)}{f_{\mathbf{x}}(x)}$	$f_{\mathbf{x}\|\mathbf{y}}(x\|y) = \frac{f_{\mathbf{y}\|\mathbf{x}}(y\|x)f_{\mathbf{x}}(x)}{f_{\mathbf{y}}(y)}$
	$P(B\|A) = \frac{P(A\|B)P(B)}{P(A)}$	$f_{\mathbf{x}\|A}(x\|A) = \frac{P(A\|\mathbf{x}=x)f_{\mathbf{x}}(x)}{P(A)}$	$f_{\mathbf{y}\|\mathbf{x}}(y\|x) = \frac{f_{\mathbf{x}\|\mathbf{y}}(x\|y)f_{\mathbf{y}}(y)}{f_{\mathbf{x}}(x)}$
TPT	$P(A) = \sum_{i=1}^{n_B} P(A\|B_i)P(B_i)$	$P(A) = \int_{-\infty}^{\infty} P(A\|\mathbf{x}=x)f_{\mathbf{x}}(x)dx$	$f_{\mathbf{y}}(y) = \int_{-\infty}^{\infty} f_{\mathbf{y}\|\mathbf{x}}(y\|x)f_{\mathbf{x}}(x)dx$
	$P(B) = \sum_{i=1}^{n_A} P(B\|A_i)P(A_i)$	$f_{\mathbf{x}}(x) = \sum_{i=1}^{n_A} f_{\mathbf{x}\|A_i}(x\|A_i)P(A_i)$	$f_{\mathbf{x}}(x) = \int_{-\infty}^{\infty} f_{\mathbf{x}\|\mathbf{y}}(x\|y)f_{\mathbf{y}}(y)dy$
Bayes Tho-erem	$P(A\|B) = \frac{P(B\|A)P(A)}{\sum_{i=1}^{n_A} P(B\|A_i)P(A_i)}$	$P(A\|\mathbf{x}=x) = \frac{f_{\mathbf{x}\|A}(x\|A)P(A)}{\sum_{i=1}^{n_A} f_{\mathbf{x}\|A_i}(x\|A_i)P(A_i)}$	$f_{\mathbf{x}\|\mathbf{y}}(x\|y) = \frac{f_{\mathbf{y}\|\mathbf{x}}(y\|x)f_{\mathbf{x}}(x)}{\int_{-\infty}^{\infty} f_{\mathbf{y}\|\mathbf{x}}(y\|x)f_{\mathbf{x}}(x)dx}$
	$P(B\|A) = \frac{P(A\|B)P(B)}{\sum_{i=1}^{n_B} P(A\|B_i)P(B_i)}$	$f_{\mathbf{x}\|A}(x\|A) = \frac{P(A\|\mathbf{x}=x)f_{\mathbf{x}}(x)}{\int_{-\infty}^{\infty} P(A\|\mathbf{x}=x)f_{\mathbf{x}}(x)dx}$	$f_{\mathbf{y}\|\mathbf{x}}(y\|x) = \frac{f_{\mathbf{x}\|\mathbf{y}}(x\|y)f_{\mathbf{y}}(y)}{\int_{-\infty}^{\infty} f_{\mathbf{x}\|\mathbf{y}}(x\|y)f_{\mathbf{y}}(y)dy}$

	Event A with Event B	Event A with RV $\mathbf{x}$	RV $\mathbf{x}$ with RV $\mathbf{y}$
Bayes Rule	$P(A\|B) = \frac{P(B\|A)P(A)}{P(B)}$ $P(B\|A) = \frac{P(A\|B)P(B)}{P(A)}$	$P(A\|\mathbf{x}=x) = \frac{P(\mathbf{x}=x\|A)P(A)}{P(\mathbf{x}=x)}$ $P(\mathbf{x}=x\|A) = \frac{P(A\|\mathbf{x}=x)P(\mathbf{x}=x)}{P(A)}$	$P(\mathbf{x}=x\|\mathbf{y}=y) = \frac{P(\mathbf{y}=y\|\mathbf{x}=x)P(\mathbf{x}=x)}{P(\mathbf{y}=y)}$ $P(\mathbf{y}=y\|\mathbf{x}=x) = \frac{P(\mathbf{x}=x\|\mathbf{y}=y)P(\mathbf{y}=y)}{P(\mathbf{x}=x)}$
TPT	$P(A) = \sum_{i=1}^{n_B} P(A\|B_i)P(B_i)$ $P(B) = \sum_{i=1}^{n_A} P(B\|A_i)P(A_i)$	$P(A) = \sum_{i=1}^{n_\mathbf{x}} P(A\|\mathbf{x}=x_i)P(\mathbf{x}=x_i)$ $P(\mathbf{x}=x) = \sum_{i=1}^{n_A} P(\mathbf{x}=x_i)\|A_i)P(A_i)$	$P(\mathbf{y}=y) = \sum_{i=1}^{n_\mathbf{x}} P(\mathbf{y}=y\|\mathbf{x}=x_i)P(\mathbf{x}=x_i)$ $P(\mathbf{x}=x) = \sum_{i=1}^{n_\mathbf{y}} P(\mathbf{x}=x\|\mathbf{y}=y_i)P(\mathbf{y}=y_i)$
Bayes Tho-erem	$P(A\|B) = \frac{P(B\|A)P(A)}{\sum_{i=1}^{n_A} P(B\|A_i)P(A_i)}$ $P(B\|A) = \frac{P(A\|B)P(B)}{\sum_{i=1}^{n_B} P(A\|B_i)P(B_i)}$	$P(A\|\mathbf{x}=x) = \frac{P(\mathbf{x}=x\|A)P(A)}{\sum_{i=1}^{n_\mathbf{y}} P(\mathbf{x}=x\|\mathbf{y}=y_i)P(\mathbf{y}=y_i)}$ $P(\mathbf{x}=x\|A) = \frac{P(A\|\mathbf{x}=x)P(\mathbf{x}=x)}{\sum_{i=1}^{n_\mathbf{x}} P(A\|\mathbf{x}=x_i)P(\mathbf{x}=x_i)}$	$P(\mathbf{x}=x\|\mathbf{y}=y) = \frac{P(\mathbf{y}=y\|\mathbf{x}=x)P(\mathbf{x}=x)}{\sum_{i=1}^{n_\mathbf{x}} P(\mathbf{y}=y\|\mathbf{x}=x_i)P(\mathbf{x}=x_i)}$ $P(\mathbf{y}=y\|\mathbf{x}=x) = \frac{P(\mathbf{x}=x\|\mathbf{y}=y)P(\mathbf{y}=y)}{\sum_{i=1}^{n_\mathbf{y}} P(\mathbf{x}=x\|\mathbf{y}=y_i)P(\mathbf{y}=y_i)}$

Plot for Example 6-42



6-7 Conditional Expected Values

Definition

Let $\mathbf{y}$ be a random variable and let $g(y)$ be a function $g : \mathbb{R} \to \mathbb{R}$. The *conditional expectation* is

$$E\{g(\mathbf{y})|M\} = \int_{-\infty}^{\infty} g(y) f_{\mathbf{y}|M}(y|M)\, dy$$

Properties

1. The *conditional mean* is

$$\mu_{\mathbf{y}|x} = E\{\mathbf{y}|x\} = \int_{-\infty}^{\infty} y f_{\mathbf{y}|x}(y|x)\, dy$$

2. The *conditional variance* is

$$\sigma^2_{\mathbf{y}|x} = E\{(\mathbf{y} - \mu_{\mathbf{y}|x})^2\} = \int_{-\infty}^{\infty} (y - \mu_{\mathbf{y}|x})^2 f_{\mathbf{y}|x}(y|x)\, dy$$

3. The conditional mean $E\{\mathbf{y}|x\}$ is a function of x.

   (a) This function of x is called a *regression line* even if the function is not a line.

   (b) Replacing x in this function by the random variable $\mathbf{x}$ creates the random variable $E\{\mathbf{y}|\mathbf{x}\}$.

   (c) The expected value of the random variable $E\{\mathbf{y}|\mathbf{x}\}$ is

$$E\{\, E\{\mathbf{y}|\mathbf{x}\}\,\} = \int_{-\infty}^{\infty} E\{\mathbf{y}|x\} f_{\mathbf{x}}(x)\, dx = E\{\mathbf{y}\}$$

4. Generalizations:

   (a) $E\{g(\mathbf{x}, \mathbf{y})|x\}$ is a function of x.

$$E\{g(\mathbf{x}, \mathbf{y})|x\} = \int_{-\infty}^{\infty} g(x, y) f_{\mathbf{y}|x}(y|x)\, dy$$

   (b) Replacing x by the random variable $\mathbf{x}$ makes $E\{g(\mathbf{x}, \mathbf{y})|\mathbf{x}\}$ a random variable.

$$E\{E\{g(\mathbf{x}, \mathbf{y})|\mathbf{x}\}\} = E\{g(\mathbf{x}, \mathbf{y})\}$$

   (c) $g(x, y) = g_1(x) g_2(y)$:

$$E\{g_1(\mathbf{x})g_2(\mathbf{y})|x\} = E\{g_1(x)g_2(\mathbf{y})|x\} = g_1(x)E\{g_2(\mathbf{y})|x\}$$
$$E\{g_1(\mathbf{x})g_2(\mathbf{y})\} = E\{E\{g_1(\mathbf{x})g_2(\mathbf{y})|\mathbf{x}\}\} = E\{g_1(\mathbf{x})E\{g_2(\mathbf{y})|\mathbf{x}\}\}$$