6.7 An area example

Let $\mathbf{x}$ and $\mathbf{y}$ be two independent random variables over the interval $[0,1]$. The density functions for $\mathbf{x}$ and $\mathbf{y}$ are given by

$$\begin{aligned} f_{\mathbf{x}}(x) &= 1 \qquad \text{if } 0 \le x \le 1 \\ &= 0 \qquad \text{otherwise;} \\ \\ f_{\mathbf{y}}(y) &= 1 \qquad \text{if } 0 \le y \le 1 \\ &= 0 \qquad \text{otherwise.} \end{aligned} \tag{7.1}$$

Because $\mathbf{x}$ and $\mathbf{y}$ are independent random variables, the joint density function $f_{\mathbf{xy}}(x,y)$ between $\mathbf{x}$ and $\mathbf{y}$ is determined by $f_{\mathbf{xy}}(x,y) = f_{\mathbf{x}}(x)f_{\mathbf{y}}(y)$. Hence

$$\begin{aligned} f_{\mathbf{xy}}(x,y) &= 1 \qquad \text{if } 0 \le x \le 1 \text{ and } 0 \le y \le 1 \\ &= 0 \qquad \text{otherwise.} \end{aligned} \tag{7.2}$$

Let $\mathbf{z}$ be the random variable given by $\mathbf{z} = \mathbf{xy}$. Notice that $\mathbf{z}$ is the area formed by $\mathbf{x}$ and $\mathbf{y}$. Our problem is to compute $E(\mathbf{x}|\mathbf{z})$. In other words, compute the best estimate of $\mathbf{x}$ given the area $\mathbf{z}$ formed by $\mathbf{x}$ and $\mathbf{y}$.

To compute $E(\mathbf{x}|\mathbf{z})$, we need the joint density function $f_{\mathbf{xz}}(x,z)$ between $\mathbf{x}$ and $\mathbf{z}$. We will use Remark 6.6.1 to compute $f_{\mathbf{xz}}(x,z)$. To this end, let

$$x = g(x,z) \quad \text{and} \quad y = h(x,z) = \frac{z}{x}.$$

In this case, the random variables, $\mathbf{u}$ and $\mathbf{y}$ are given by

$$\mathbf{u} = g(\mathbf{x},\mathbf{z}) = \mathbf{x} \quad \text{and} \quad \mathbf{y} = h(\mathbf{x},\mathbf{z}) = \frac{\mathbf{z}}{\mathbf{x}}.$$

The mapping $[g(x,z), h(x,z)]^{tr} \mapsto [x, \frac{z}{x}]^{tr}$ is almost everywhere an invertible mapping from $\mathbb{R}^2$ onto $\mathbb{R}^2$. In this case, the Jacobian is determined by

$$J = \begin{bmatrix} \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial z} \\ \\ \dfrac{\partial h}{\partial x} & \dfrac{\partial h}{\partial z} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\dfrac{z}{x^2} & \dfrac{1}{x} \end{bmatrix}.$$

Notice that $\det[J] = 1/x$. Recall that $f_{\mathbf{xy}}(x,y) = f_{\mathbf{x}}(x)f_{\mathbf{y}}(y)$. By virtue of Remark 6.6.1, the joint density function $f_{\mathbf{xz}}(x,z)$ is given by

$$f_{\mathbf{xz}}(x,z) = f_{\mathbf{xy}}(g(x,z),h(x,z))|\det J| = f_{\mathbf{x}}(g(x,y))f_{\mathbf{y}}(h(x,z))\frac{1}{x} = f_{\mathbf{x}}(x)f_{\mathbf{y}}(z/x)\frac{1}{x}.$$

Since $f_{\mathbf{y}}(y)$ has support in $[0,1]$, the function $f_{\mathbf{y}}(\frac{z}{x}) \neq 0$ if and only if $0 \le \frac{z}{x} \le 1$, or equivalently, $0 \le z \le x$. Since $f_{\mathbf{x}}(x)$ has support in $[0,1]$, the variable x must be in $[0,1]$. Combining this with $0 \le z \le x$, we see that $f_{\mathbf{x}}(x)f_{\mathbf{y}}(z/x)$ is nonzero in $0 \le z \le x$ and $0 \le x \le 1$. By consulting (7.1) or (7.2), we obtain

$$\begin{aligned} f_{\mathbf{xz}}(x,z) &= \frac{1}{x} \qquad \text{if } \ 0 < z \le x \le 1 \\ &= \qquad \text{otherwise.} \end{aligned} \tag{7.3}$$

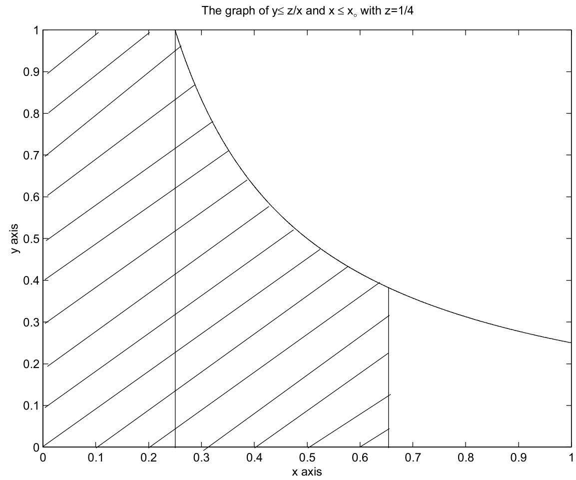

Figure 6.3: The area $\mathcal{A}$

A quadratic approximation of $\widehat{g}(\mathbf{y})$.

Another method to compute $f_{\mathbf{xz}}(x,z)$. Recall that for fixed $x_\circ$ and z, the distribution function $F_{\mathbf{xz}}(x_\circ,z) = P(\mathbf{x} \le x_\circ \bigcap \mathbf{z} \le z)$. Using $\mathbf{xy} = \mathbf{z}$, we obtain

$$\begin{aligned} F_{\mathbf{xz}}(x_\circ,z) &= P(\mathbf{x} \le x_\circ \bigcap \mathbf{z} \le z) = P(\mathbf{x} \le x_\circ \bigcap \mathbf{xy} \le z) \\ &= P([\mathbf{x},\mathbf{y}] \in \mathcal{A}) = \int\int_{\mathcal{A}} f_{\mathbf{xy}}(x,y)\,dydx. \end{aligned}$$

Here the area $\mathcal{A}$ in $\mathbb{R}^2$ is determined by the set of all $[x,y]$ in $\mathbb{R}^2$ such that $0 \le x \le 1$ and $0 \le y \le 1$ and $x \le x_\circ \le 1$ and $y \le \frac{z}{x}$. This area is below the graph of $y = \frac{z}{x}$, to the left of the line $x = x_\circ$ and contained in the square $0 \le x \le 1$ and $0 \le y \le 1$. Because the support for $f_{\mathbf{xy}}(x,y)$ is contained in the square $0 \le x \le 1$ and $0 \le y \le 1$, the area $\mathcal{A}$ must also be contained in this box. Since $0 \le y \le 1$ and $y = \frac{z}{x}$, we obtain that $0 \le \frac{z}{x} \le 1$. In other words, $0 \le z \le x \le 1$. Finally, observe that for $y = 1$, we have $z = x$. A typical graph of this region $\mathcal{A}$ is given by the shaded region in Figure 6.3. The first vertical line corresponds to $x = z$, while the second vertical line corresponds to $x = x_\circ$. So we see that

$$\begin{aligned} F_{\mathbf{xz}}(x_\circ,z) &= \int\int_{\mathcal{A}} f_{\mathbf{xy}}(x,y)\,dydx = \int_0^z\int_0^1 dydx + \int_z^{x_\circ}\int_0^{\frac{z}{x}} dydx \\ &= z + \int_z^{x_\circ} \frac{z}{x}dx = z + z\ln(x_\circ) - z\ln(z). \end{aligned}$$

By replacing $x_\circ$ with x, we obtain

$$F_{\mathbf{xz}}(x,z) = z + z\ln(x) - z\ln(z) \qquad \text{if } 0 \le z \le x \le 1.$$

Using the fact that

$$f_{\mathbf{xz}}(x,z) = \frac{\partial^2}{\partial x \partial z} F_{\mathbf{xz}}(x,z),$$

we arrive at

$$\begin{aligned} f_{\mathbf{xz}}(x,z) &= \frac{1}{x} && \text{if } \ 0 < z \le x \le 1 \\ &= && \text{otherwise.} \end{aligned} \tag{7.4}$$

The density function $f_{\mathbf{z}}(z)$ for the random variable $\mathbf{z}$ is computed by

$$\begin{aligned} f_{\mathbf{z}}(z) = \int_{-\infty}^{\infty} f_{\mathbf{xz}}(x,z)dx &= \int_z^1 \frac{1}{x}dx = -\ln(z) && \text{if } 0 \le z \le 1 \\ &= \int_{-\infty}^{\infty} 0dx = 0 && \text{otherwise.} \end{aligned}$$

In other words, the density function $f_{\mathbf{z}}(z)$ for $\mathbf{z}$ is given by

$$\begin{aligned} f_{\mathbf{z}}(z) &= -\ln(z) && \text{if } 0 \le z \le 1 \\ &= 0 && \text{otherwise.} \end{aligned} \tag{7.5}$$

Finally, it is noted that $f_{\mathbf{z}}(z)$ is indeed positive and has area one. Since $-\ln(z) \ge 0$ for all $0 < z \le 1$, it follow that $f_{\mathbf{z}}(z) \ge 0$ for all z. Using the change of variable $z = e^{\lambda}$, we obtain

$$\int_{-\infty}^{\infty} f_{\mathbf{z}}(z)dz = -\int_0^1 \ln(z)dz = -\int_{-\infty}^0 \lambda e^{\lambda} d\lambda = 1.$$

As expected, $f_{\mathbf{z}}(z)$ is positive and has area one.

The conditional density $f_{\mathbf{x}|\mathbf{z}}(x|z)$ is only defined for $0 < z \le 1$. By consulting (7.3) and (7.5), the conditional density $f_{\mathbf{x}|\mathbf{z}}(x|z)$ is given by

$$\begin{aligned} f_{\mathbf{x}|\mathbf{z}}(x|z) = \frac{f_{\mathbf{xz}}(x,z)}{f_{\mathbf{z}}(z)} &= \frac{-1}{x\ln(z)} && \text{if } 0 < z \le x \le 1 \\ &= 0 && \text{otherwise.} \end{aligned}$$

Notice that for any fixed z in the interval $[0,1]$, the conditional density $f_{\mathbf{x}|\mathbf{z}}(x|z)$ is a density function in x, that is, $f_{\mathbf{x}|\mathbf{z}}(x|z) \ge 0$ and $1 = \int_{-\infty}^{\infty} f_{\mathbf{x}|\mathbf{z}}(x|z)dx$. The optimal function

$$\widehat{g}(z) = E(\mathbf{x}|\mathbf{z} = z) = \int_{-\infty}^{\infty} x f_{\mathbf{x}|\mathbf{z}}(x|z)dx = \int_z^1 \frac{-1}{\ln(z)}dx = \frac{z-1}{\ln(z)}.$$

Therefore the optimal estimate is given by

$$\widehat{g}(z) = E(\mathbf{x}|\mathbf{z} = z) = \frac{z-1}{\ln(z)} \qquad (0 \le z \le 1). \tag{7.6}$$

For example, if the area $z = \frac{1}{10}$, then the best estimate of x is given by $\widehat{g}(\frac{1}{10}) \approx 0.39$. If the area $z = \frac{1}{100}$, then the best estimate of x is given by $\widehat{g}(\frac{1}{100}) \approx 0.215$.

Recall that the best estimate $\widehat{g}(\mathbf{z}) = E(\mathbf{x}|\mathbf{z})$ is the random variable which is the unique solution to the optimization problem,

$$E|\mathbf{x} - \widehat{g}(\mathbf{z})|^2 = \inf\{E|\mathbf{x} - g(\mathbf{z})|^2 : g \text{ is a measurable function from } \mathbb{R}^{\nu} \text{ into } \mathbb{C}\}. \tag{7.7}$$

In this case, the optimal solution is given by

$$\widehat{g}(\mathbf{z}) = E(\mathbf{x}|\mathbf{z}) = \frac{\mathbf{z}-1}{\ln(\mathbf{z})}. \tag{7.8}$$

Recall that $E\widehat{g}(\mathbf{z}) = EE(\mathbf{x}|\mathbf{z}) = E\mathbf{x}$. Since $\mathbf{x}$ is uniform over $[0,1]$, we see that $E\mathbf{x} = 1/2$. Hence $E\widehat{g}(\mathbf{z}) = 1/2$. To verify this directly, simply observe that

$$E\widehat{g}(\mathbf{z}) = E\left(\frac{\mathbf{z}-1}{\ln(\mathbf{z})}\right) = \int_{-\infty}^{\infty} \frac{z-1}{\ln(z)} f_{\mathbf{z}}(z)dz = \int_0^1 (1-z)dz = \frac{1}{2}.$$

The covariance error in estimation is computed by

$$\begin{aligned} E|\mathbf{x}-\widehat{g}(\mathbf{z})|^2 &= E\mathbf{x}^2 - E\widehat{g}(\mathbf{z})^2 = \int_{-\infty}^{\infty} x^2 f_{\mathbf{x}}(x)dx - \int_{-\infty}^{\infty} \widehat{g}(z)^2 f_{\mathbf{z}}(z)dz \\ &= \int_0^1 x^2 dx + \int_0^1 \frac{(z-1)^2}{\ln(z)}dz = \frac{1}{3} + \ln(3/4) \approx 0.0457. \end{aligned}$$

In other words, the error in estimation is given by

$$\|\mathbf{x}-\widehat{g}(\mathbf{z})\| = \sqrt{E|\mathbf{x}-\widehat{g}(\mathbf{z})|^2} \approx 0.2138. \tag{7.9}$$

A third order polynomial approximation of $\widehat{g}(\mathbf{z})$. Now let us obtain an approximation of the best estimate $\widehat{g}(\mathbf{z})$ using polynomials of order at most three. To this end, consider the vector

$$\mathbf{p} = \begin{bmatrix} 1 \\ \mathbf{z} \\ \mathbf{z}^2 \\ \mathbf{z}^3 \end{bmatrix}.$$

Let $\mathcal{M}$ be the four dimensional space spanned by $\mathbf{p}$, that is,

$$\mathcal{M} = \operatorname{span}\{\mathbf{p}\} = \operatorname{span}\{1, \mathbf{z}, \mathbf{z}^2, \mathbf{z}^3\}.$$

Recall that the orthogonal projection $\widehat{\mathbf{x}} = P_{\mathcal{M}}\mathbf{x}$ is computed by

$$\begin{aligned} \widehat{\mathbf{x}} &= P_{\mathcal{M}}\mathbf{x} = R_{\mathbf{xp}}R_{\mathbf{p}}^{-1}\mathbf{p} \\ E(\mathbf{x}-\widehat{\mathbf{x}})^2 &= R_{\mathbf{x}} - R_{\mathbf{xp}}R_{\mathbf{p}}^{-1}R_{\mathbf{px}}. \end{aligned} \tag{7.10}$$

Notice that

$$\begin{aligned} R_{\mathbf{xp}} &= E\mathbf{xp}^* = \begin{bmatrix} E\mathbf{x} & E\mathbf{xz} & E\mathbf{xz}^2 & E\mathbf{xz}^3 \end{bmatrix} \\ R_{\mathbf{p}} &= E\mathbf{pp}^* = \begin{bmatrix} E1 & E\mathbf{z} & E\mathbf{z}^2 & E\mathbf{z}^3 \\ E\mathbf{z} & E\mathbf{z}^2 & E\mathbf{z}^3 & E\mathbf{z}^4 \\ E\mathbf{z}^2 & E\mathbf{z}^3 & E\mathbf{z}^4 & E\mathbf{z}^5 \\ E\mathbf{z}^3 & E\mathbf{z}^4 & E\mathbf{z}^5 & E\mathbf{z}^6 \end{bmatrix}. \end{aligned}$$

Recall that $\mathbf{x}$ and $\mathbf{y}$ are two independent uniform random variables over the interval $[0,1]$. For any positive integer k, we have

$$
\begin{aligned}
E\mathbf{x}\mathbf{z}^k &= E\mathbf{x}^{k+1}\mathbf{y}^k = E\mathbf{x}^{k+1}E\mathbf{y}^k = \int_0^1 x^{k+1}dx \times \int_0^1 y^k dy = \frac{1}{(k+2)(k+1)} \\
E\mathbf{z}^k &= E\mathbf{x}^k\mathbf{y}^k = E\mathbf{x}^k E\mathbf{y}^k = \int_0^1 x^k dx \times \int_0^1 y^k dy = \frac{1}{(k+1)^2} \\
R_{\mathbf{x}} &= E\mathbf{x}^2 = \int_0^1 x^2 dx = \frac{1}{3}.
\end{aligned}
$$

This readily implies that

$$
\begin{aligned}
R_{\mathbf{xp}} &= \begin{bmatrix} \frac{1}{2} & \frac{1}{6} & \frac{1}{12} & \frac{1}{20} \end{bmatrix} \\
R_{\mathbf{p}} &= \begin{bmatrix} 1 & \frac{1}{4} & \frac{1}{9} & \frac{1}{16} \\ \frac{1}{4} & \frac{1}{9} & \frac{1}{16} & \frac{1}{25} \\ \frac{1}{9} & \frac{1}{16} & \frac{1}{25} & \frac{5}{36} \\ \frac{1}{16} & \frac{1}{25} & \frac{5}{36} & \frac{1}{49} \end{bmatrix}.
\end{aligned}
$$

Using Matlab we see that

$$
\begin{aligned}
\widehat{\mathbf{x}} &= R_{\mathbf{xp}}R_{\mathbf{p}}^{-1}\mathbf{p} = 0.2140 + 1.8416\mathbf{z} - 2.3412\mathbf{z}^2 + 1.3717\mathbf{z}^3 \\
E(\mathbf{x} - \widehat{\mathbf{x}})^2 &= R_{\mathbf{x}} - R_{\mathbf{xp}}R_{\mathbf{p}}^{-1}R_{\mathbf{px}} \approx 0.0459. \qquad (7.11)
\end{aligned}
$$

The error in estimation is given by

$$
\|\mathbf{x} - \widehat{\mathbf{x}}\| = \sqrt{E(\mathbf{x} - \widehat{\mathbf{x}})^2} \approx 0.2143. \qquad (7.12)
$$

As expected, the error from the conditional expectation is smaller than the error from the polynomial approximation, that is,

$$
\sqrt{E(\mathbf{x} - \widehat{g}(\mathbf{z}))^2} \approx 0.2138 \leq 0.2143 \approx \sqrt{E(\mathbf{x} - \widehat{\mathbf{x}})^2}.
$$

The graph given in Figure 6.4 plots the optimal function

$$
\widehat{g}(z) = \frac{z-1}{\ln(z)}
$$

and its best approximation $0.2140 + 1.8416z - 2.3412z^2 + 1.3717z^3$ by polynomials of degree at most three on the same graph. As expected both of these plots are close to each other. The graph corresponding to $\widehat{g}(z)$ is zero at $z = 0$ and has a smaller value at $z = 1$.

Monte Carlo. Let us use the Monte Carlo technique to estimate $\widehat{g}(\mathbf{z}) = E(\mathbf{x}|\mathbf{z} = z)$. Ruffly speaking, the Monte Carlo method uses a computer simulation of the appropriate random values to compute the expectation or conditional expectation for a random variable. This is particularly useful when the conditional expectation is difficult to compute. Let us

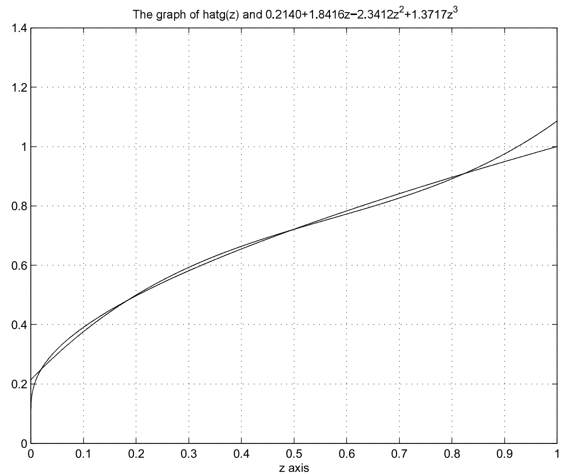

Figure 6.4: The conditional expectation $\widehat{g}(z) = \frac{z-1}{\ln(z)}$

use Matlab to compute the conditional expectation $E(\mathbf{x}|\mathbf{z})$ evaluated at $\mathbf{z} = \frac{1}{2}$, that is, let us use the Monte Carlo method to compute an approximation of

$$\widehat{g}(1/2) = E(\mathbf{x}|\mathbf{z} = 1/2) = \left.\frac{z-1}{\ln(z)}\right|_{z=1/2} = -\frac{1}{2\ln(1/2)} = 0.7213.$$

To this end, first use Matlab to compute $x = \text{rand}(1, 50000)$; and $y = \text{rand}(1, 50000)$; which are 50000 samples of the uniform random variables $\mathbf{x}$ and $\mathbf{y}$ over the interval $[0, 1]$, respectively. Then $z = x.*y$ in Matlab is 50000 samples of the random variable $\mathbf{z}$. According to Matlab the mean of $\mathbf{z}$ computed from these samples is given by

$$\text{mean}(z) = 0.2494 \approx E\mathbf{z} = E\mathbf{x}E\mathbf{y} = \frac{1}{4}.$$

To estimate $\widehat{g}(1/2)$, we compute the mean of all $x(k)$ such that $z(k) = x(k) * y(k) \approx \frac{1}{2}$. (Notice that numerically $z(k)$ will rarely if ever equal $\frac{1}{2}$. So we take the mean of all $x(k)$ such that $z(k)$ is in some neighborhood of $\frac{1}{2}$.) In Matlab we used the commands:

$$c = [\,]; \text{for } k = 1 : 50000; \text{if abs}(x(k) * y(k) - 1/2) < .01, c = [c, x(k)]; \text{end;end}$$

Then we discovered that $\text{mean}(c) = 0.7222 \approx 0.7213 = \widehat{g}(1/2)$.

Exercise 1. Let $\mathbf{x}$ and $\mathbf{y}$ be two continuous independent random variables. Let $\mathbf{z}$ be the random variable defined by $\mathbf{z} = \mathbf{xy}$. Then show that

$$f_{\mathbf{xz}}(x, z) = f_{\mathbf{x}}(x) f_{\mathbf{y}}(z/x) \frac{1}{|x|}. \tag{7.13}$$

Moreover, the density for $\mathbf{z}$ is given by

$$f_{\mathbf{z}}(z) = \int_{-\infty}^{\infty} f_{\mathbf{x}}(x) f_{\mathbf{y}}(z/x) \frac{dx}{|x|}. \tag{7.14}$$