1. Recall that we already know that $Z \sim P(\lambda + \mu)$, right?

Given that $Z = n \geq 0$, the value of $X$ can only be one of $k = 0, 1, \ldots, n$. By the definition of conditional probability (given an event), for such values of $k$, one has

$$
P(X = k|Z = n) = \frac{P(X = k, Z = n)}{P(Z = n)} = \frac{P(X = k, Y = n - k)}{P(Z = n)}
$$

$$= \frac{P(X = k)P(Y = n - k)}{P(Z = n)} \quad \text{[by independence]}
$$

$$= \frac{e^{-\lambda} \frac{\lambda^k}{k!} \times e^{-\mu} \frac{\mu^{n-k}}{(n-k)!}}{e^{-\lambda - \mu} \frac{(\lambda + \mu)^n}{n!}} = \binom{n}{k} \frac{\mu^k (1 - p)^{n-k}},$$

which is the binomial distribution $B(n, p)$ with $p = \frac{\lambda}{\lambda + \mu}$.

2. (a) In this case, given that $Y(\omega) = y$, one must have $\omega = \pm \sqrt{y}$, so $Y$ “partitions” $[-1, 1]$ into two-point sets of the form $\{-\omega, \omega\}$, the “atoms” of $\sigma(Y)$. Now $\hat{X}$ is obtained by averaging $X$ over such “atoms”, and as the probability $P$ is uniform on $[-1, 1]$, one would expect to have $\hat{X}(\omega) = \frac{1}{2}(X(\omega) - X(-\omega))$. Now let’s show that formally!

For $B = (a, b), 0 < a < b < 1$, one has $\{Y \in B\} = (-\sqrt{b}, -\sqrt{a}) \cup (\sqrt{a}, \sqrt{b})$. By definition, $\hat{X}(\omega) = h(Y) = h(\omega^2)$, so that $\hat{X}(\omega) = \hat{X}(\omega)$, and, as $P (d\omega) = \frac{1}{2} d\omega$ on $[-1, 1]$ ($P$ has a density $\frac{1}{2}$ on our $\Omega = [-1, 1]$, right?),

$$E(\hat{X}; Y \in B) = \int_{-\sqrt{b}}^{-\sqrt{a}} \hat{X}(\omega) \frac{1}{2} d\omega + \int_{\sqrt{a}}^{\sqrt{b}} \hat{X}(\omega) \frac{1}{2} d\omega
$$

$$= \int_{\sqrt{a}}^{\sqrt{b}} \frac{1}{2} (\hat{X}(\omega) + \hat{X}(\omega)) d\omega = \int_{\sqrt{a}}^{\sqrt{b}} \hat{X}(\omega) d\omega.
$$

On the other hand, using the same argument,

$$E(X; Y \in B) = \int_{\sqrt{a}}^{\sqrt{b}} \frac{1}{2} (X(\omega) + X(\omega)) d\omega = \int_{\sqrt{a}}^{\sqrt{b}} X(\omega) d\omega.
$$

Clearly, $X_0(\omega) = X_0(-\omega)$, so that $X_0$ is a function of $|\omega|$ or, equivalently, of $\omega^2 = Y$, and, as such, it has the same integrals as $\hat{X}$ over all subintervals $B$ of $[0, 1]$. That can be extended to all Borel subsets of $[0, 1]$ (it’s the same as extending a probability measure from intervals), which means that $P (X_0 = \hat{X}) = 1$. So we showed that

$$\hat{X} = E(X|Y) = \frac{1}{2} (X(-\omega) + X(\omega)) = \frac{1}{2} (X(\sqrt{Y}) + X(-\sqrt{Y})).
$$

(b) Here $Y(\omega) = \omega^3$ is a one-to-one function on $[-1, 1]$. One has $\omega = Y^{1/3}$ and hence $X(\omega) = X(Y^{1/3})$ is already a function of $Y$. By (CEP.3), $E(X|Y) = X$.  

1
(c) Part (a): we still have $\hat{X}(\omega) = \hat{X}(-\omega)$, but now (again for $B = (a, b)$, $0 < a < b < 1$) one has

$$E(X; Y \in B) = \int_{-\sqrt{b}}^{\sqrt{a}} \hat{X}(\omega) f(\omega) d\omega + \int_{\sqrt{a}}^{\sqrt{b}} \hat{X}(\omega) f(\omega) d\omega$$

$$= \int_{\sqrt{a}}^{\sqrt{b}} (\hat{X}(-\omega) f(-\omega) + \hat{X}(\omega) f(\omega)) d\omega = \int_{\sqrt{a}}^{\sqrt{b}} \hat{X}(\omega) (f(-\omega) + f(\omega)) d\omega,$$

and similarly

$$E(X; Y \in B) = \int_{\sqrt{a}}^{\sqrt{b}} (X(-\omega) f(-\omega) + X(\omega) f(\omega)) d\omega.$$

So one must have $\hat{X}(\omega)(f(-\omega) + f(\omega)) = X(-\omega)f(-\omega) + X(\omega)f(\omega)$, yielding

$$\hat{X}(\omega) = \frac{X(-\omega)f(-\omega) + X(\omega)f(\omega)}{f(-\omega) + f(\omega)} \quad \text{a.s.}$$

Part (b): the same story as in (b). Regardless of the underlying probability, $X(\omega) = X(Y^{1/3})$ is already a function of $Y$. So, by (CEP.3), $E(X|Y) = X$.

3. (a) Assume for simplicity that $f$ is continuous. Look: for $x \in \mathbb{R}$ and small $\Delta > 0$,

$$f_X(x) \Delta \approx P(X \in (x, x + \Delta)) = P(X \in (x, x + \Delta), Y \in \mathbb{R})$$

$$= \int \int_{(x+x+\Delta) \times \mathbb{R}} f(u, y) du \ dy = \int \left[ \int_{x}^{x+\Delta} f(u, y) du \right] dy$$

$$\approx \Delta \times \int f(x, y) \ dy,$$

where we used our continuity assumption to replace $f(u, y)$ inside the narrow strip $(x, x + \Delta) \times \mathbb{R}$ with $f(x, y)$. This kind of shows that $f_X(x) = \int f(x, y) \ dy$. Well, the above is not a rigorous proof, but it should be OK for our purposes.

(b) To compute the joint density $f$, use the general formula (*) from slide 89. Since the covariance matrix has the form

$$C^2 = \begin{bmatrix} \sigma_X^2 & \rho \sigma_X \sigma_Y \\ \rho \sigma_X \sigma_Y & \sigma_Y^2 \end{bmatrix}, \quad \text{with} \quad \det C^2 = (1 - \rho) \sigma_X^2 \sigma_Y^2,$$

the inverse matrix equals (using the standard minors/cofactors’ method)

$$[C^2]^{-1} = \frac{1}{\det C^2} \begin{bmatrix} \sigma_Y^2 & -\rho \sigma_X \sigma_Y \\ -\rho \sigma_X \sigma_Y & \sigma_X^2 \end{bmatrix} = \frac{1}{1 - \rho^2} \begin{bmatrix} \frac{1}{\sigma_X^2} & -\frac{\rho}{\sigma_X \sigma_Y} \\ -\frac{\rho}{\sigma_X \sigma_Y} & \frac{1}{\sigma_Y^2} \end{bmatrix}.$$ 

Now, putting for brevity $x_0 := (x - \mu_X)/\sigma_X$ and $y_0 := (y - \mu_Y)/\sigma_Y$, one has

$$(x - \mu_X, y - \mu_Y) \begin{bmatrix} \frac{1}{\sigma_X^2} & -\frac{\rho}{\sigma_X \sigma_Y} \\ -\frac{\rho}{\sigma_X \sigma_Y} & \frac{1}{\sigma_Y^2} \end{bmatrix} (x - \mu_X, y - \mu_Y) = x_0^2 - 2\rho x_0 y_0 + y_0^2.$$
and hence, from (*),

\[ f(x, y) = \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1 - \rho^2}} \exp\left\{-\frac{x_0^2}{2(1 - \rho^2)} - \frac{2\rho x_0 y_0 + y_0^2}{2(1 - \rho^2)} \right\}, \]

where the quadratic form in the numerator of the fraction in the exponential can be re-written as \((y_0 - \rho x_0)^2 + (1 - \rho^2)x_0^2\).

Using the last representation for the quadratic form, we compute \(f_X\) by integrating the joint density w.r.t. \(y\):

\[
f_X(x) = \int f_{(X,Y)}(x,y)dy \]

\[
= \frac{e^{-x_0^2/2}}{\sqrt{2\pi \sigma_X}} \cdot \frac{1}{\sqrt{2\pi \sigma_Y \sqrt{1 - \rho^2}}} \int \exp\left\{-\frac{(y_0 - \rho x_0)^2}{2(1 - \rho^2)} \right\} \sigma_Y dy \\
= \frac{e^{-x_0^2/2}}{\sqrt{2\pi \sigma_X}} \cdot \frac{1}{\sqrt{2\pi \sigma_Y \sqrt{1 - \rho^2}}} \exp\left\{-\frac{(x - \mu_X)^2}{2\sigma_X^2} \right\}.
\]

Dividing the expression for \(f\) by \(f_X(x)\) yields the conditional density \(f_{Y|X}(y|x)\) in the form

\[
\frac{1}{\sqrt{2\pi \sigma_Y \sqrt{1 - \rho^2}}} \exp\left\{-\frac{(y_0 - \rho x_0)^2 + (1 - \rho^2)x_0^2}{2(1 - \rho^2)} + \frac{x_0^2}{2} \right\} \\
= \frac{1}{\sqrt{2\pi \sigma_Y \sqrt{1 - \rho^2}}} \exp\left\{-\frac{(y_0 - \rho x_0)^2}{2(1 - \rho^2)} \right\} \\
= \frac{1}{\sqrt{2\pi \sigma_Y \sqrt{1 - \rho^2}}} \exp\left\{-\frac{(y - \mu_Y - \rho \sigma_Y (x - \mu_X)/\sigma_X)^2}{2(1 - \rho^2)\sigma_Y^2} \right\},
\]

so that the desired condition distribution is

\[ N(\mu_Y + \rho \frac{\sigma_Y}{\sigma_X} (x - \mu_X), (1 - \rho^2)\sigma_Y^2). \]

(c) This will be just the expectation of the conditional distribution, with \(x\) replaced by \(X\): \(E(Y | X) = \mu_Y + \rho \frac{\sigma_Y}{\sigma_X} (X - \mu_X)\).

4. One has \(f_{X|Z}(x|z) = f_{(X,Z)}(x,z)/f_Z(z)\), where, by convolution formula, for \(z > 0\),

\[
f_Z(z) = f_{X+Y}(z) = \int f_X(x)f_Y(z-x)dx \\
= \int \lambda e^{-\lambda x}1(x > 0)\lambda e^{-\lambda(z-x)}1(z-x > 0)dx = \int_0^z x^2 e^{-\lambda x}dx = \lambda^2ze^{-\lambda z}.
\]

To find \(f_{(X,Z)}\), note that \((X,Z) = (X,X+Y) = (X,Y) \begin{pmatrix}1 & 1 \\ 0 & 1 \end{pmatrix} =: (X,Y)A\) is a linear function \(g\) of \((X,Y)\), with the inverse \(h(x,z) = (x,z)A^{-1} = (x,z-x)\) (observe that \(A^{-1} = \begin{pmatrix}1 & -1 \\ 0 & 1 \end{pmatrix}\)), with Jacobian \(Jh \equiv 1\). As \(f_{(X,Y)}(x,y) = f_X(x)f_Y(y) = \lambda e^{-\lambda x}1(x > 0)\lambda e^{-\lambda y}1(y > 0)\) by independence, by Thm. [2.43] one has

\[
f_{(X,Z)}(x,z) = f_{(X,Y)}(h(x,z)) = \lambda e^{-\lambda x}1(x > 0)\lambda e^{-\lambda(z-x)}1(z-x > 0) = \lambda^2e^{-\lambda z}1(0 < x < z).
\]
Hence the desired conditional density equals
\[ f_{X|Z}(x|z) = \frac{x^2 e^{-\lambda x} \mathbf{1}(0 < x < z)}{\lambda^2 z e^{-\lambda x} \mathbf{1}(z > 0)} = \frac{1}{z} \mathbf{1}(z > 0), \]

which is the density of the uniform distribution \( U(0, z) \).

5. (a) The RV \( X \) is discrete, taking values \( k = 0, 1, 2, \ldots \), and, by the total probability formula,\(^1\)

\[
P(X = k) = \mathbb{E} P(X = k | Y) = \int P(X = k | Y = y) \, dF_Y(y) = \frac{1}{k!} \int_0^\infty y^k e^{-y} dy = \frac{1}{2k+1} \int_0^\infty u^k e^{-u} du = 2^{-k-1}.
\]

(b) Using the definition of conditional probability/expectation, for \( B = [0, x] \) for any fixed \( x > 0 \), one has, using the result of part (a), for \( k = 0, 1, 2, \ldots \),

\[
P(Y \leq x | X = k) = \mathbb{P}(Y \in B | X = k) = \frac{\mathbb{P}(Y \in B, X = k)}{\mathbb{P}(X = k)} = 2^{k+1} \mathbb{E} (1(X = k); Y \in B)
\]

\[= 2^{k+1} \mathbb{E} \left[ \mathbb{E} (1(X = k) | Y); Y \in B \right] = 2^{k+1} \int_0^x \frac{y^k}{k!} e^{-y} \times e^{-y} dy.
\]

That is, the conditional distribution of \( Y \) given \( X = k \) is \( \gamma(k + 1, 2) \).

Using mathematical “shorthand”, the above derivation can be written more concisely (and intuitively) as

\[
P(Y \in dy | X = k) = \frac{\mathbb{P}(Y \in dy, X = k)}{\mathbb{P}(X = k)} = 2^{k+1} \mathbb{P}(X = k | Y = y) \mathbb{P}(Y \in dy)
\]

\[= 2^{k+1} \times \frac{y^k}{k!} e^{-y} \times y^{-y} dy = \frac{2^{k+1}}{k!} y^ke^{-2y} dy, \quad y > 0.
\]

Density plotting: DIY. One has, for \( y > 0 \), the following functions:

\[
k = 0 : \quad 2e^{-2y}; \quad k = 1 : \quad 4ye^{-2y}; \quad k = 2 : \quad 8y^2 e^{-2y}.
\]

(c) Since \( \mathbb{E}(X|Y = y) = \text{the expectation of } P(y) \), which is \( y \), one has \( \mathbb{E}(X|Y) = Y \).

Now \( \mathbb{E}(Y|X = k) = \text{the expectation of } \gamma(k + 1, 2) \), which is \( (k + 1)/2 \). Indeed, the expectation of \( \gamma(a, b) \), is

\[
\gamma(a, b) = \int_0^\infty y \frac{b^a}{\Gamma(a)} y^{a-1} e^{-by} dy = \frac{\Gamma(a + 1)}{b\Gamma(a)} \int_0^\infty \frac{b^{a+1}}{\Gamma(a+1)} y^ae^{-by} dy = \frac{a}{b}, \quad = 1, \text{ the integral of gamma density}
\]

\(^1\)Look: if \( Y \) has density \( f \), then \( \mathbb{P}(A) = \mathbb{E} 1_A = \mathbb{E} \mathbb{E}(1_A|Y) = \mathbb{E} \mathbb{P}(A|Y) = \int \mathbb{P}(A|Y = y) f(y) dy \), right?
So $E(Y|X) = (X + 1)/2$.

Finally,

$$\text{Cov}(X, Y) = E XY - E X E Y \overset{(\text{Cov.5})}{=} E E(XY|Y) - \left(E \left(E(X|Y)\right)\right)E Y$$

$$= E \left(Y \left(E(X|Y)\right)\right) - (E Y)^2 = E Y^2 - (E Y)^2 = \text{Var}(Y) = 1, \quad \text{right?}$$