Problem VI.1: Let \((X,Y)\) be the coordinates of a point picked uniformly in a circle centered at the origin and with radius 1. Calculate
\begin{enumerate}
  \item \(E[X \mid Y \geq a]\) for \(a \in (-1, +1)\);
  \item \(E(|X - Y|^2)\);
  \item \(E[X \mid X \geq Y]\);
  \item \(E[X \mid X + Y]\).
\end{enumerate}

a. Recall that \(X\) is a real valued function over the sample space - i.e. \(X : \Omega \mapsto \mathbb{R}\). When we condition on an event that occurs we change the probability distribution of events in the sample space. In this way the statistics of the random variable \(X\) change. But the random variable \(X\) itself maps values from the sample space to the real line just as before. \(X\)'s randomness is completely inherited from the randomness of its arguments. So \(X\)'s distribution changes with the probability assignments of its arguments. Thus \(E[X \mid Y \geq a]\) is just a short hand for compute the average value of \(X\) using the probability assignment of events after conditioning on the event \(\{Y \geq a\}\). We can express this more formally as
\[
\frac{E[X1(Y \geq a)]}{P(Y \geq a)} \tag{1}
\]

Note that this is nothing more than the definition of conditional probability in action. The indicator controls the values of \(X\) we average over - namely we average over only those values of \(X\) that occur when the event \(\{Y \geq a\}\) occurs. Thus the numerator expresses nothing other than an intersection of two events - the subset of values of \(X\) which have non-zero probability of occurring when \(\{T \geq a\}\) occurs. The denominator renormalizes our computation because probability must be redistributed over the new sample space such that the probability over it is 1. Proceeding we get:
\[
E[X \mid Y \geq a] = \frac{E[X1(Y \geq a)]}{P(Y \geq a)} = \frac{\int_{-1}^{1} \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} x1(y \geq a) \frac{1}{\pi} dx dy}{P(Y \geq a)}
\]
\[
= \frac{\int_{-1}^{1} 1(y \geq a) \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} x dx dy}{P(Y \geq a)}
\]
\[
= \frac{\int_{a}^{1} \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} x dx dy}{P(Y \geq a)}
\]
\[
= \frac{\int_{a}^{1} \frac{1}{2}[(1-y^2) - (1-y^2)] dy}{P(Y \geq a)}
\]
\[
= 0 \tag{2}
\]

b. One way to approach this problem is to work in polar coordinates.
\[
E[|X - Y|^2] = E[|R \cos \Theta - R \sin \Theta|^2] \\
= E[R^2 | \cos \Theta - \sin \Theta|^2] \\
= \int_0^1 \int_0^{2\pi} r^2 | \cos \theta - \sin \theta|^2 \frac{1}{\pi} r \theta dr d\theta \\
= \frac{1}{\pi} \int_0^1 r^3 dr \int_0^{2\pi} | \cos \theta - \sin \theta|^2 d\theta \\
= \frac{1}{4\pi} \int_0^{2\pi} 2 \cos^2(\theta + \phi) d\theta \\
= \frac{1}{4\pi} \int_0^{2\pi} 2 \cos^2 \theta d\theta \\
= \frac{1}{4\pi} \cdot 2\pi \\
= \frac{1}{2} \\
\]

Note that \(\phi\) really doesn't matter since we are integrating over a positive integer multiple of the period - namely over an interval that is twice as long as the period which is \(\pi\).

c. First note that the conditioning event means that the point \((X, Y)\) is now uniformly distributed over the half circle \((R, \Theta) \in [0, 1] \times [-\frac{\pi}{4}, \frac{\pi}{4}]\). Thus,

\[
E[X \mid X \geq Y] = \frac{E[X \mathbb{1}(X \geq Y)]}{P(X \geq Y)} \\
= \frac{E[R \cos \Theta \mathbb{1}(\cos \Theta \geq \sin \Theta)]}{\frac{1}{2}} \\
= 2 \int_0^1 \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} r \cos \theta \frac{1}{\pi} r \theta dr d\theta \\
= \frac{2}{\pi} \int_0^1 r^2 dr \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \cos \theta d\theta \\
= \frac{2}{\pi} \left( \frac{1}{3} \right) \left( \sin \left( \frac{\pi}{4} \right) + \sin \left( \frac{3\pi}{4} \right) \right) \\
= \frac{2\sqrt{2}}{3\pi} \\
\]

Here's a long way to do the problem. Solve the problem \(E[Y \mid Y \geq 0]\) and take the projection of this value on the line \(Y = X\). Note that

\[
f_Y(y \mid Y \geq 0) = \frac{4\sqrt{1 - y^2}}{\pi} \quad (5)
\]

This can be obtained by determining \(F_Y(y \mid Y \geq 0) = \frac{F_Y(y)}{P(Y \geq 0)}\) and differentiating the result. I'll let you do that. Using the density function \(f_Y(y \mid Y \geq 0)\) we can integrate to obtain the expectation.
\[ E[Y|Y \geq 0] = \int_0^1 y f_Y(y \mid Y \geq 0) dy \]
\[ = \int_0^1 y \frac{4\sqrt{1-y^2}}{\pi} dy \]
\[ = \frac{4}{\pi} \int_0^1 \cos^2 \theta \sin \theta d\theta \]
\[ = \frac{4}{\pi} \left( \frac{1}{3} \right) \int_0^1 u^2 du \]
\[ = \frac{4}{3\pi} \]

To obtain the projection we scale this answer by \( \cos\left(\frac{\pi}{4}\right) \).

\[ E[X|X \geq Y] = \frac{\sqrt{2}}{2} \frac{4}{3\pi} \]
\[ = \frac{2\sqrt{2}}{3\pi} \]

The quick way to do this problem is by a symmetry argument. By symmetry \( E[X \mid X+Y] = E[Y \mid X+Y] \). Note also that \( E[X+Y \mid X+Y] = X+Y \). But by linearity of conditional expectation we get \( E[X+Y \mid X+Y] = E[X \mid X+Y] + E[Y \mid X+Y] \). Thus \( E[X \mid X+Y] = \frac{X+Y}{2} \).

We can go through a more tedious approach using the distribution of \((X,Y)\) conditioned on \(X+Y\). Simply note if we observe that \(X+Y = u\) then the point \((X,Y)\) is uniformly drawn from the segment that is the intersection of the line \(y = -x+u\) and the unit disk. Thus, conditioned on \(X+Y\), \(X\) is uniformly distributed over \(\left[\frac{1}{2}(X+Y-\sqrt{2}-(X+Y)^2), \frac{1}{2}(X+Y+\sqrt{2}-(X+Y)^2)\right]\). The end points of this interval are nothing other than the x-coordinate of the two points at which the line \(y = -x+u\) intersects the unit disk. Thus the mean value of \(X\) conditioned on \(X+Y\) is just \(\frac{1}{2}(X+Y)\).

**Problem VI.2:** Assume that \(X\) and \(Y\) are independent and uniformly distributed in \([0, 1]\). Calculate \(E[X|X^2 + Y^2] \).

Note that \(X^2 + Y^2 \in [0, 2]\). The expectation can be projected onto two distinct cases. In the first case \(X^2 + Y^2 \leq 1\). In the second case \(X^2 + Y^2 > 1\). Note that using polar representation given \(X^2 + Y^2\), \(X = \sqrt{X^2 + Y^2} \cos \Theta\) where \(\Theta \in [0, \frac{\pi}{2}]\) uniformly. Thus

\[ E[X|X^2 + Y^2, X^2 + Y^2 \leq 1] = \int_0^{\frac{\pi}{2}} \sqrt{X^2 + Y^2} \cos \Theta \frac{2}{\pi} d\Theta \]
\[ = \frac{2}{\pi} \sqrt{X^2 + Y^2} \]

The second case is a bit trickier. Again we resort to a polar representation of the point \((X,Y)\). Note that \(\Theta\) is now uniform in \([\phi, \frac{\pi}{2} - \phi]\) where \(\phi\) is given by

\[ \phi = \arctan \sqrt{X^2 + Y^2 - 1} \]

Thus we compute the expected value of \(X\) in a similar way to the first case.
$$E[X|X^2 + Y^2, X^2 + Y^2 > 1] = \int_{\phi}^{\pi - \phi} \sqrt{X^2 + Y^2} \cos \theta \frac{1}{\pi - 2\phi} \, d\theta$$

$$= \frac{2}{\pi - 4\phi} \sqrt{X^2 + Y^2} (\sin(\frac{\pi}{2} - \phi) - \sin \phi)$$

$$= \frac{2}{\pi - 4\phi} \sqrt{X^2 + Y^2} (\cos \phi - \sin \phi)$$

$$= \frac{2}{\pi - 4\phi} \sqrt{X^2 + Y^2} \left( \sqrt{\frac{1}{X^2 + Y^2}} - \sqrt{\frac{X^2 + Y^2 - 1}{X^2 + Y^2}} \right)$$

$$= \frac{2}{\pi - 4 \arctan(\sqrt{X^2 + Y^2 - 1})} (1 - \sqrt{X^2 + Y^2 - 1})$$

Putting it all together

$$E[X|X^2 + Y^2] = E[X|X^2 + Y^2, X^2 + Y^2 \leq 1]1(X^2 + Y^2 \leq 1)$$

$$+ E[X|X^2 + Y^2, X^2 + Y^2 > 1]1(X^2 + Y^2 > 1)$$

$$= \frac{2}{\pi} \sqrt{X^2 + Y^2} 1(X^2 + Y^2 \leq 1)$$

$$+ \frac{2(1 - \sqrt{X^2 + Y^2 - 1})}{\pi - 4 \arctan(\sqrt{X^2 + Y^2 - 1})} 1(X^2 + Y^2 > 1)$$

Problem VI.3: Let \((X, Y)\) designate the coordinates of a point picked uniformly on the circumference of a circle with radius one centered at the origin. Calculate

$$E[X + Y|X + Y > 0].$$

Let \((1, \Theta)\) denote the polar representation of the point \((X, Y)\) where \(\Theta\) is the angle made with respect to the positive \(x\)-axis. Note that \(\Theta\) is distributed uniformly on the interval \([-\pi, \pi]\). By conditioning on the event \(\{X + Y > 0\}\), \(\Theta\) is distributed uniformly over \((-\frac{\pi}{4}, \frac{3\pi}{4})\). Note that \(X = \cos \Theta\) and \(Y = \sin \Theta\).

$$E[X|X + Y > 0] = E[\cos \Theta|\Theta \in (-\frac{1}{4}\pi, \frac{3}{4}\pi)]$$

$$= \frac{E[\cos \Theta|\Theta \in (-\frac{1}{4}\pi, \frac{3}{4}\pi)]}{Pr(\Theta \in (-\frac{1}{4}\pi, \frac{3}{4}\pi))}$$

$$= \int_{-\frac{1}{4}\pi}^{\frac{1}{4}\pi} \cos \theta \frac{1}{2\pi} \, d\theta$$

$$= \frac{1}{\pi} \left[ \int_{-\frac{1}{4}\pi}^{\frac{1}{4}\pi} \cos \theta \, d\theta \right]$$

$$= \frac{1}{\pi} \left[ \sin(\frac{3\pi}{4}) + \sin(\frac{\pi}{4}) \right]$$

$$= \frac{\sqrt{2}}{\pi}$$
We obtain similar equations for $Y$.

\[
E[Y|X+Y > 0] = E[\sin \Theta | \Theta \in \left( -\frac{1}{4}\pi, \frac{3}{4}\pi \right)] \\
= \int_{-\frac{1}{4}\pi}^{\frac{1}{4}\pi} \sin \theta \frac{1}{\pi} d\theta \\
= \frac{1}{\pi} \left( \cos \left( \frac{\pi}{4} \right) - \cos \left( \frac{3\pi}{4} \right) \right) \\
= \frac{\sqrt{2}}{\pi}
\]  

(13)

Thus $E[X + Y|X + Y > 0] = \frac{2\sqrt{2}}{\pi}$.

**Problem VI.4:** Let $X, Y, Z$ be independent and exponentially distributed with mean 1. Calculate

\[
E[X|X+Y+Z].
\]


\[
X + Y + Z = E[X + Y + Z|X + Y + Z] \\
= 3E[X|X + Y + Z]
\]  

(14)

Thus,

\[
E[X|X + Y + Z] = \frac{1}{3}(X + Y + Z)
\]  

(15)

**Problem VI.5:** Assume that $X$ and $Y$ are independent and uniformly distributed in $[0, 1]$. Calculate

\[
E[(X + Y)^3|X].
\]

\[
E[(X + Y)^3|X] = E[X^3 + 3X^2Y + 3XY^2 + Y^3|X] \\
= E[X^3|X] + 3E[X^2Y|X] + 3E[XY^2|X] + E[Y^3|X] \\
= X^3 + 3X^2E[Y|X] + 3XE[Y^2|X] + E[Y^3|X] \\
= X^3 + 3X^2 \left( \frac{1}{2} \right) + 3X \left( \frac{1}{3} \right) + \frac{1}{4} \\
= X^3 + \frac{3}{2}X^2 + X + \frac{1}{4}
\]  

(16)

**Problem VI.6:** Assume that an event $A$ happens with an unknown probability $p$ in a single trial. The value of $p$ is picked uniformly in $[0, 1]$. Find the expected value of $p$ given that the event $A$ occurs
a times out of $a + b$ independent replications of the trial.

Apply this result to the following example. You flip a coin that has an unknown probability of yielding H. The first $n$ flips yield H. What is the probability that the next flip will also yield H?

Note: When $a$ and $b$ are integers,

$$
\int_0^1 x^{a-1}(1 - x)^{b-1} dx = \frac{(a - 1)!(b - 1)!}{(a + b - 1)!}.
$$

Let $A_i$ denote the event that on the $i^{th}$ trial event $A$ occurs. Let $X_{a+b} = \sum_{i=1}^{a+b} 1(A_i)$. $X_{a+b}$ counts the number of times $A$ occurs in $a + b$ trials. Conditioned on $p$, $X_{a+b}$ is distributed binomial($a + b, p$). Bayes’ Rule gives us the following:

$$
P(p = u \mid X_{a+b} = a) = \frac{P(X_{a+b} = a \mid p = u)P(p = u)}{P(X_{a+b} = a)}
= \frac{(a+b)^a}{a! b!} \int_0^1 u^a (1 - u)^b du
= \frac{(a + b + 1)!(a + 1)!}{a! b!(a + b + 2)!}
$$

Note that I am use $P(p = u \mid X_{a+b} = a)$ as shorthand for $f_p(u \mid X_{a+b} = a)$ and $P(p = u)$ for $f_p(u)$. The basic premise behind the equality is Bayes’ Rule, but it doesn’t follow immediately from it since Bayes’ Rule deals only with probabilities not densities.

To find the conditional expectation we integrate $u$ with respect to $P(p = u \mid X_{a+b} = a)du$.

$$
E[p \mid X_{a+b} = a] = \int_0^1 uP(p = u \mid X_{a+b} = a)du
= \frac{(a + b + 1)!}{a! b!} \int_0^1 u^{a+1} (1 - u)^b du
= \frac{(a + b + 1)!}{a! b!} \frac{(a + 1)!}{(a + b + 2)!}
= \frac{a + 1}{a + b + 2}
$$

For the coin tossing example, the event that the next flip also yields a H given that the first $n$ yield H is the same as the event $\{X_{n+1} = n + 1\}$ conditioned on the event $\{X_n = n\}$. 
\[ P(X_{n+1} = n + 1 \mid X_n = n) = \frac{P(\{X_{n+1} = n + 1\} \cap \{X_n = n\})}{P(\{X_n = n\})} \]
\[ = \frac{P(\{X_{n+1} = n + 1\})}{P(\{X_n = n\})} \]
\[ = \frac{\int_0^1 P(\{X_{n+1} = n + 1\} \mid p = u) f_p(u) du}{\int_0^1 P(\{X_n = n\} \mid p = u) f_p(u) du} \]
\[ = \frac{\int_0^1 P(\{X_{n+1} = n + 1\} \mid p = u) du}{\int_0^1 P(\{X_n = n\} \mid p = u) du} \]
\[ = \frac{\int_0^1 \binom{n+1}{n+1} u^{n+1} (1 - u)^0 du}{\int_0^1 \binom{n}{n} u^n (1 - u)^0 du} \]
\[ = \frac{\int_0^1 u^{n+1} du}{\int_0^1 u^n du} \]
\[ = \frac{n + 1}{n + 2} \quad (19) \]

Note that this can be obtained from the previous result by plugging in for \(a = n\) and \(b = 0\) since the probability that the next flip is H just the conditional expectation of \(p\) given you observed a string of \(n\) H.

**Problem VI.7:** Let \(X\) and \(Y\) be independent and uniformly distributed in \([0, 1]\). Calculate \(E[X\mid XY]\).

One way to approach this problem is to use the conditional density:

\[ f_{V\mid U}(v\mid U) = \frac{f_{V,U}(U, v)}{f_U(U)} \quad (20) \]

where \(U = XY\) and \(V = X\). We can obtain \(f_{V,U}(v, u)\) from \(f_{X,Y}(x, y)\) using the general change of variables formula.

\[ f_{V,U}(v, u) = f_{X,Y}(x(v, u), y(v, u))|J(v, u)| \quad (21) \]

where \(x(v, u) = u\) and \(y(v, u) = \frac{u}{v}\) and \(J(v, u)\) is given by

\[ J(v, u) = \begin{vmatrix} \frac{\partial x(v, u)}{\partial v} & \frac{\partial y(v, u)}{\partial v} \\ \frac{\partial x(v, u)}{\partial u} & \frac{\partial y(v, u)}{\partial u} \end{vmatrix} = \begin{vmatrix} 0 & 1/v \\ 1 & -u/v^2 \end{vmatrix} = -v^{-1} \quad (22) \]

Recall that

\[ f_{X,Y}(x, y) = 1(0 \leq x \leq 1)1(0 \leq y \leq 1) \quad (23) \]

Thus
\[ f_{VU}(v, u) = f_{XY}(v, \frac{u}{v})v^{-1} \]
\[ = 1(0 \leq v \leq 1)1(0 \leq \frac{u}{v} \leq 1)v^{-1} \]

We need \( f_U(u) \). To get it we average out \( V \).
\[
f_U(u) = \int_{-\infty}^{\infty} f_{VU}(v, u)dv \]
\[ = \int_{0}^{1} 1(0 \leq \frac{u}{v} \leq 1)v^{-1}dv \]
\[ = \int_{0}^{1} v^{-1}dv \]
\[ = \log(1) - \log(u) \]
\[ = -\log(u) \]

Thus
\[
f_{V|U}(v \mid U) = \frac{1(0 \leq v \leq 1)1(0 \leq \frac{U}{v} \leq 1)}{-v \log(U)} \]

Finally we evaluate \( E[X \mid XY] = E[V \mid U] \)
\[
E[V \mid U] = \int_{-\infty}^{\infty} vf_{V|U}(v \mid U)dv \\
= \int_{-\infty}^{\infty} v\frac{1(0 \leq v \leq 1)1(0 \leq \frac{U}{v} \leq 1)}{-v \log(U)}dv \\
= \int_{0}^{1} \frac{1}{v - \log(U)}dv \\
= \frac{1 - U}{\log\left(\frac{1}{U}\right)} \\
= \frac{1 - XY}{\log\left(\frac{1}{XY}\right)}
\]

Problem VI.8: Let \( X \) be a random variable with mean \( \mu \) and finite variance \( \sigma^2 \). Show that the value of \( a \) that minimizes \( E([X - a]^2) \) is \( a = \mu \).

Note the following equality for an arbitrary real \( a \) when \( X \) has finite variance. (If \( X \) does not have finite variance it makes no sense to write down the second moments of a random variable since they do not exist.)
\[
E[(X - a)^2] = E[(X - \mu + \mu - a)^2]
\]
\[
= E[(X - \mu)^2] + E[2(X - \mu)(\mu - a)] + E[(\mu - a)^2]
\]
\[
= E[(X - \mu)^2] + 2(\mu - a)E[(X - \mu)] + (\mu - a)^2
\]
\[
= E[(X - \mu)^2] + (\mu - a)^2
\]  
(28)

Note however that \((\mu - a)^2 \geq 0\) with equality only when \(a = \mu\). Adding \(E[(X - \mu)^2]\) to both sides of this inequality gives

\[
E[(X - a)^2] \geq E[(X - \mu)^2]
\]  
(29)

Thus \(E[(X - a)^2]\) for arbitrary real \(a\) is bounded below by \(E[(X - \mu)^2]\). Equality holds if \(a = \mu\). Thus the mean square difference is minimized when \(a = \mu\). This means that the mean of a random variable \(X\) is the optimal point estimator of \(X\) with respect to the mean square error metric.

**Problem VI.9:** Given \(Y\), let \(X_1, X_2, \ldots\) be independent and uniformly distributed on \([0, Y]\).

a. Assume that \(Y\) is uniformly distributed on \([0, A]\). Find \(E[Y | \max\{X_1, \ldots, X_n\}]\).

b. Repeat the problem, assuming that \(Y\) is exponentially distributed with mean \(A\).

a. One approach is to use the conditional density \(f_{Y | \max\{X_1, \ldots, X_n\}}(u | \max\{X_1, \ldots, X_n\})\). Consider first the case where \(n > 2\).

\[
E[Y | \max\{X_1, \ldots, X_n\}] = \int_0^A u f_{Y | \max\{X_1, \ldots, X_n\}}(u | \max\{X_1, \ldots, X_n\})du
\]  
(30)

Note that we use Bayes’ rule for densities to obtain the desired density. Let \(Z = \max\{X_1, \ldots, X_n\}\).

\[
f_{Y|Z}(u | Z) = \frac{f_{Z|Y}(Z|Y = u)f_Y(u)}{f_Z(Z)}
\]  
(31)

We first find the joint density using the distribution of \(Z\) conditioned on \(Y\).

\[
F_{Z|Y}(z | Y = u) = P(Z \leq z | Y = u)
\]
\[
= P(\max\{X_1, \ldots, X_n\} \leq z | Y = u)
\]
\[
= P(X_1 \leq z, \ldots, X_n \leq z | Y = u)
\]
\[
= \prod_{i=1}^n P(X_i \leq z | Y = u)
\]
\[
= P(X_1 \leq z | Y = u)^n
\]
\[
= (\frac{z}{u})^n 1(0 \leq z \leq u) + 1(z > u)
\]  
(32)

Differentiating the result with respect to \(z\):

\[
f_{Z|Y}(z | Y = u) = n \frac{z^{n-1}}{u^n} 1(0 \leq z \leq u)
\]  
(33)
Note that

$$f_Y(u) = \frac{1}{A} 1(0 \leq u \leq A) \quad (34)$$

To find the marginal density of $Z$ we integrate the joint density over all values $Y$ takes on.

$$f_Z(z) = \int_{-\infty}^{\infty} f_{Z|Y}(z|u)f_Y(u)du$$

$$= \int_{-\infty}^{\infty} n\frac{z^{n-1}}{u^n} 1(0 \leq z \leq u) \frac{1}{A} 1(0 \leq u \leq A)du$$

$$= n \frac{z^{n-1}}{A} \int_{z}^{A} u^{-n}du$$

$$= n \frac{z^{n-1}}{A(n-1)} (u^{1-n} - A^{1-n})$$

$$= \frac{n}{A(n-1)} (1 - \left(\frac{z}{A}\right)^{n-1}) \quad (35)$$

Finally we compute the conditional expectation:

$$E[Y \mid Z] = \int_{-\infty}^{\infty} u \frac{Z^{n-1}}{u^n} 1(0 \leq Z \leq u) \frac{1}{A} 1(0 \leq u \leq A)du$$

$$= (n-1) \frac{Z^{n-1}}{A(n-1)} \int_{Z}^{A} u^{-n}du$$

$$= \frac{n-1}{n-2} \frac{Z - \frac{1}{A}(Z)^{n-1}}{1 - \left(\frac{Z}{A}\right)^{n-1}}$$

$$= \frac{n-1}{n-2} \frac{\max\{X_1, \ldots, X_n\} - \frac{1}{A}(\max\{X_1, \ldots, X_n\})^{n-1}}{(1 - \left(\frac{\max\{X_1, \ldots, X_n\}}{A}\right)^{n-1})} \quad (36)$$

Note that $\lim_{n \to \infty} E[Y \mid \max\{X_1, \ldots, X_n\}] = \max\{X_1, \ldots, X_n\}$. This matches our intuition because with each additional $X_i$ the $Z$ can only get closer to $Y$. We can actually make the strong statement that the sequence of random variables $Z_i = \max\{X_1, \ldots, X_n\}$ converges with probability one - or converges almost surely - to the random variable $Y$. This means that any realization of a sequence of $Z_i(\omega)$ will converge to $Y(\omega)$ eventually. Recall that random variables are functions over the sample space. When we say converge we mean pointwise convergence - not uniform convergence. We can also prove that the sequence of conditional expectations also converge almost surely to the random variable $Y$. This is good news. The formalism and models match our intuition. We will talk more about what it means for a sequence of random variables to converge to another random variable later in the course.

Note the above does not hold when $n \leq 2$. If you go through the same methodology you will find that:

$$E[Y \mid \max\{X_1\}] = \frac{A - \max\{X_1\}}{\log \frac{A}{\max\{X_1\}}} \quad (37)$$

$$E[Y \mid \max\{X_1, X_2\}] = \frac{A \max\{X_1, X_2\} \log \frac{A}{\max\{X_1, X_2\}}}{A - \max\{X_1, X_2\}}$$
b. Again I opted to take the conditional expectation approach. All the relevant pieces to making the computation were obtained using the same steps as in part a. Let $Z$ be defined as before. Then

$$f_{Z|Y}(z|Y = u) = n \frac{z^{n-1}}{u^n} 1(0 \leq z \leq u)$$

$$f_Y(u) = A e^{-Au} 1(u \geq 0)$$

$$f_Z(z) = \int_z^{\infty} n \frac{z^{n-1}}{u^n} A e^{-Au} du$$

Putting everything together

$$E[Y|Z] = \frac{\int_z^{\infty} u^{1-n} e^{-Au} du}{\int_z^{\infty} u^{-n} e^{-Au} du}$$

The antiderivatives of the functions under the integrals involve a function called a partial Gamma function. I’ll leave those interested in partial Gamma functions to simplify the integrals on their own. :-)

\text{---}