4. Consider a sample of positive random variables $X_1, X_2, \ldots, X_n$.

   a. Show that the arithmetic mean is greater than or equal to the geometric mean, which is in turn greater than or equal to the harmonic mean.

   b. Under what conditions will exact equality hold between any two of the above descriptive statistics?

   c. Show that the median of the sample can be in any relation to the three means.

   Answer:

   This problem is easily solved making use of Jensen's inequality.

   We need the following results about convex functions:

- Definition (Monotonicity; Convexity) A function $g(x)$ is said to be
  a. *monotonically nondecreasing* if for all $a < b$ in the domain of g, $g(a) \leq g(b)$.
  b. *monotonically increasing* if for all $a < b$ in the domain of g, $g(a) < g(b)$.
  c. *monotonically nonincreasing* if for all $a < b$ in the domain of g, $g(a) \geq g(b)$.
  d. *monotonically decreasing* if for all $a < b$ in the domain of g, $g(a) > g(b)$.
  e. *monotonic* if it is either monotonically nonincreasing or monotonically nondecreasing.
  f. *strictly monotonic* if it is either monotonically increasing or monotonically decreasing.
  g. *convex* if for all $a < b$ in the domain of g and all $p \in (0,1)$, $g(pa + (1-p)b) \leq pg(a) + (1-p)g(b)$.
  h. *strictly convex* if for all $a < b$ in the domain of g and all $p \in (0,1)$, $g(pa + (1-p)b) < pg(a) + (1-p)g(b)$.
  i. *concave* if $-g(x)$ is convex.
  j. *strictly concave* if $-g(x)$ is strictly convex.
- Proposition A function $g(x)$ that is twice differentiable is convex if $g''(x) \geq 0$ for all x, and it is strictly convex if $g''(x) > 0$ for all x.
- Proposition (Jensen's Inequality) Let $g(x)$ be a convex function. Then for random variable X, $E[g(X)] \geq g(E[X])$. If $g(X)$ is strictly convex, then $E[g(X)] > g(E[X])$.
  - (Note: The direction of the inequality is easy to remember by the following: $g(x) = x^2$ is convex, and because variances must be nonegative, $E[X^2] - E^2[X] \geq 0$.)

Now for the answers:

a. We want first to show that the geometric mean is less than the arithmetic mean

$$\exp\left(\frac{1}{n}\sum_{i=1}^{n}\log(X_i)\right) \leq \frac{1}{n}\sum_{i=1}^{n} X_i.$$

We can re-write that equation as

$$\exp\left(\frac{1}{n}\sum_{i=1}^{n}\log(X_i)\right) \leq \frac{1}{n}\sum_{i=1}^{n}\exp\left(\log(X_i)\right),$$

which is of the form

$$g(E[Y]) \leq E[g(Y)]$$

for $g(y) = \exp(y)$ and $Y = \log(X)$. Now $g(y)$ is easily shown to be strictly convex, because $g''(y) = \exp(y) > 0$ for all y. Hence, Jensen's inequality gives us our result.

Now we want to show the harmonic mean is less than the geometric mean

$$\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}\right)^{-1} \leq \exp\left(\frac{1}{n}\sum_{i=1}^{n}\log(X_i)\right).$$

The above equation will be true whenever

$$-\log\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}\right) \leq \left(\frac{1}{n}\sum_{i=1}^{n}\log(X_i)\right),$$

which can in turn be re-written as

$$-\log\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}\right) \leq \left(\frac{1}{n}\sum_{i=1}^{n}-\log\left(\frac{1}{X_i}\right)\right).$$

This is now of the form

$$g(E[Y]) \leq E[g(Y)]$$

for $g(y) = -\log(y)$ and $Y = (1/X)$. Now $g(y)$ is easily shown to be strictly convex, because $g''(y) = 1/X^2 > 0$ for all y. Hence, Jensen's inequality gives us our result.

b. Note that in both of the above applications of Jensen's inequality, the functions were strictly convex. Hence the only way that equality can hold is if $X_i = X_j$ for all i and j. That is, exact equality holds only for degenerate distributions–constants.

c. We will find distributions in which the median is less than the harmonic mean, the median is between the harmonic mean and the geometric mean, the median is between the geometric mean and the arithmetic mean, or the median is greater than the arithmetic mean.

- Let $X \in \{1, 2\}$ with $Pr(X = 2) = 0.75$ and $Pr(X = 2) = 0.25$. Then the harmonic mean is 8/7 and the median is 1.
- Let $X \in \{0.5, 1, 2\}$ with $Pr(X = 0.5) = 0.2$, $Pr(X = 1) = 0.5$, and $Pr(X = 2) = 0.3$. Then the harmonic mean is 0.95, the median is 1, and the geometric mean is 1.07.
- Let $X \in \{0.5, 1, 2\}$ with $Pr(X = 0.5) = 0.49$, $Pr(X = 1) = 0.03$, and $Pr(X = 2) = 0.48$. Then the geometric mean is 0.993, the median is 1, and the arithmetic mean is 1.235.
- Let $X \in \{1, 2\}$ with $Pr(X = 1) = 0.25$ and $Pr(X = 2) = 0.75$. Then the arithmetic mean is 1.75 and the median is 2.