Grading: 1, 2, 3, 4, 5 (each 8 pts).

Problem 1: **a)** Note that $E[X^p] = \int_{-\infty}^{\infty} x^p d\mu(x) = \int_1^\infty x^p r x^{-r-1} dx$ so is finite iff $p-r-1 < -1$, i.e. $p < r$ and for that case $E[X^p] = r/x^{r-p}(p-r)|_1^\infty = r/(r-p)$. **b)** We define $F(x) = 0$ for $x \le e$ and $F(x) = 1 - 1/\log(x)$ for $x > e$. This has density function $f(x) = 0$ for $x < e$ and $f(x) = 1/(x \log^2(x))$ for $x > e$. Then $E[X^p] = \int_e^\infty x^{p-1}/\log^2(x) dx$. Substitute $u = \log x$ to get $\int_1^\infty e^{pu}/u^2 du$. The integrand tends to infinity as $u \to \infty$ (since exponential growth dominates polynomial growth, or by applying the L'Hopital's Rule twice) so the integral diverges.

Problem 2: **a)** Note that $P(A \cap B^c) = P(A) - P(B \cap A) = P(A) - P(A)P(B) = P(A)P(B^c)$, where the second to last equality holds if A and B are independent. Similarly, we can get that A^c and B or A^c and B^c are independent. Now let $\mathbf{1}_A$ and $\mathbf{1}_B$ be independent, then $P(A \cap B) = P(\mathbf{1}_A = 1, \mathbf{1}_B = 1) = P(\mathbf{1}_A = 1)P(\mathbf{1}_B = 1) = P(A)P(B)$ so A and B are independent. Conversely, if A and B are independent, then note that for $C, D \in \mathcal{B}$, $\{\mathbf{1}_A \in C\} \in \{\emptyset, A, A^c, \Omega\}$ and same for $\mathbf{1}_B : \{\mathbf{1}_B \in D\} \in \{\emptyset, B, B^c, \Omega\}$. If $\{\mathbf{1}_A \in C\} = A$ or A^c and $\{\mathbf{1}_B \in D\} = B$ or B^c, then by the previous argument they are independent. If $\{\mathbf{1}_A \in C\} = \emptyset$ or $\{\mathbf{1}_B \in D\} = \emptyset$ then their intersection is also the empty set, so the indicator function are again independent. The last case when $\{\mathbf{1}_A \in C\} = \Omega$ and $\{\mathbf{1}_B \in C\} = \Omega$ is analogous. **b).** Take any $A \in \sigma(X)$, $B \in \sigma(Y)$, then $A = \{X \in C\}$, $B = \{Y \in D\}$ for some $C, D \in \mathcal{B}$ and since X, Y are independent, $P(A \cap B) = P(\{X \in C\} \cap \{Y \in D\}) = P(\{X \in C\})P(\{Y \in D\}) = P(A)P(B)$ so A and B are independent. But, since they were arbitrary in $\sigma(X)$, $\sigma(Y)$, then by definition, $\sigma(X)$, $\sigma(Y)$ are independent. For the opposite, if $X \in \mathcal{F}_X$ and $Y \in \mathcal{F}_Y$, then for any $C, D \in \mathcal{B}$, by measurability $\{X \in C\} \in \mathcal{F}_X$ and $\{Y \in D\} \in \mathcal{F}_Y$. Hence, because $\mathcal{F}_X, \mathcal{F}_Y$ are independent $P(\{X \in C\} \cap \{Y \in D\}) = P(\{X \in C\})P(\{Y \in D\})$ and since C, D were arbitrary we get that X and Y are independent.

Problem 3: We define X_1, X_2 and X_3 be independent random variables taking only values -1 and 1 with equal probability. Next, define $X_4 = X_1 X_2 X_3$ and since $1/16 = P(X_1 = -1)^4 \neq 1/8 = P(X_1 = -1, X_2 = -1, X_3 = -1, X_4 = -1)$, all four are not independent. But, for $i < j < 4$, $c_1, c_2, c_3 \in \{-1, 1\}$ and if we denote $k = \{1,2,3\} - \{i,j\}$, $P(X_i = c_1, X_j = c_2, X_4 = c_3) = P(X_i = c_1, X_j = c_2, X_i X_j X_k = c_3) = P(X_i = c_1, X_j = c_2, X_k = c_3/c_1c_2) = P(X_i = c_1)P(X_j = c_2)P(X_k = c_3/c_1c_2) = 1/8$ for any of the cases, from where we can see that any three are independent.

Problem 4: After denoting an independent copy of X by Y, note that since f and g are non-decreasing, $(f(x)-f(y))(g(x)-g(y)) \ge 0, \forall x, y \in \mathbb{R}$ and hence $E[f(X)-f(Y))(g(X)-g(Y)] \ge 0 \Leftrightarrow E[f(X)g(X)] - E[f(Y)g(X)] - E[f(X)g(Y)] + E[f(Y)g(Y)] \ge 0$. (By using the Cauchy-Schwarz inequality and the fact that $f, g \in L^2$, it is easy to see that each of these four expectations is finite.) Since X and Y are copies of each other, $E(f(X)g(X)) = E(f(Y)g(Y))$ and besides, $f, g \in L^2$ implies $f, g \in L^1$, which means we can use the Theorem 6.4 to get $E(f(X)g(Y)) = E(f(X))E(g(Y)) = E(f(X))E(g(X))$. Plugging in the last two equations in the inequality we get $2E[f(X)g(X)] - 2E[f(X)]E[g(X)] \geq 0$, which completes the proof.

Problem 5: By the linearity of expectation, for $m \leq n$ we have that $E[S_m/S_n] = E[(\sum_{i=1}^{m} X_i)/S_n] = \sum_{i=1}^{m} E[X_i/S_n] = mE[X_1/S_n]$ and by setting $m = n$ we hence get $1 = nE[X_1/S_n]$, meaning that $E[S_m/S_n] = m/n$. For $m > n$, we have $E[S_m/S_n] = E[1 + (X_{n+1} + ... + X_m)/S_n] = 1 + \sum_{i=n+1}^{m} E[X_i/S_n] = 1 + (m-n)E(X_1)E[1/S_n]$, as desired. *(The last equality is true by the Theorem 6.4 combined with Corollary 6.2 of the lecture notes.)*