Problem 9.1. Recall the definitions of $x^+$ and $x^-$ given in the middle of page 405. Observe that the function 
$$f(x) \equiv x^+ \equiv \max\{0, x\}$$
is a nondecreasing function of $x$. Similarly, the function 
$$h(x) \equiv x^- \equiv -\min\{0, x\}$$
is a nonincreasing function of $x$. The arguments below apply to any such monotone functions. We only prove part (a). In (b) the inequality switches because the function is nonincreasing.

Proof 1. Apply Proposition 9.1.2 on page 405. Observe that for any nondecreasing real-valued function $g$ on $\mathbb{R}^+$, $(g \circ f)(x) \equiv g(f(x))$ is a nondecreasing real-valued function on $\mathbb{R}$. By Proposition 9.1.2, since $X \geq_{st} Y$,
$$E[(g \circ f)(X)] \geq E[(g \circ f)(Y)] ,$$so that
$$E[g(X^+)] = E[(g \circ f)(X)] \geq E[(g \circ f)(Y)] \geq E[g(Y^+)] ,$$which implies that
$$X^+ \geq_{st} Y^+ .$$

Proof 2. Apply Proposition 9.2.2 on page 405. As above, observe that the function 
$$f(x) \equiv x^+ \equiv \max\{0, x\}$$
is a nondecreasing function of $x$. Since $X \geq_{st} Y$, there exist random variables $\tilde{X}$ and $\tilde{Y}$, where
$$P(\tilde{X} \geq \tilde{Y}) = 1 ,$$
$\tilde{X}$ is distributed the same as $X$ and $\tilde{Y}$ is distributed the same as $Y$. Then
$$P(\tilde{X}^+ \geq \tilde{Y}^+) = 1 .$$
Since $\tilde{X}$ is distributed the same as $X$ and $\tilde{Y}$ is distributed the same as $Y$, $\tilde{X}^+$ is distributed the same as $X^+$ and $\tilde{Y}^+$ is distributed the same as $Y^+$. As a consequence,
$$\tilde{X}^+ \geq_{st} \tilde{Y}^+ .$$

Proof 3. Apply the definition on page 404. Note that, for $a > 0$,
$$P(X^+ > a) = P(X > a) \geq P(Y > a) = P(Y^+ > a) .$$
Problem 9.2. The ordering is true if there is independence; i.e., if \( X_1 \) is independent of \( X_2 \) and if \( Y_1 \) is independent of \( Y_2 \). To see that, use the w.p.1 coupling construction (Proposition 9.2.2).

To get the requested counterexample, relax the independence condition. An easy way is to assume that the \( X_i \) are totally dependent, while the \( Y_i \) are independent. That is let \( X_1 = X_2 \) with probability one. For simplicity, let all four random variables be Bernoulli with the same distribution as \( X_1 \):

\[
P(X_1 = 1) = 1 - P(X_1 = 0) = 1/2.
\]

Then

\[
P(X_1 + X_2 = 2) = 1/2 > 1/4 = P(Y_1 + Y_2 = 2),
\]

while

\[
P(X_1 + X_2 = 0) = 1/2 > 1/4 = P(Y_1 + Y_2 = 0).
\]

So there is no stochastic order between \( X_1 + X_2 \) and \( Y_1 + Y_2 \).

Problem 9.3. Suppose that \( X \) and \( Y \) are independent. Let \( Y_1 \) be a random variable distributed as \( Y \) that is independent of \( X \) and \( Y \). First note that

\[
P(Y_1 \geq Y) = P(Y \geq Y_1) \geq 1/2.
\]

We might have strict inequality because we may have

\[
P(Y_1 = Y) > 0.
\]

Then

\[
P(X \geq Y) = \int P(X \geq Y|Y = y)dP(Y = y)
\]

\[
= \int P(X \geq y)dP(Y = y)
\]

\[
\geq \int P(Y_1 \geq y)dP(Y = y)
\]

\[
\geq P(Y_1 \geq Y) \geq 1/2.
\]

Problem 9.4 Suppose that

\[
P_1 \geq P_2 \geq \cdots \geq P_n.
\]

Then the optimal order is \( 1, 2, \ldots, n \). The probability that the requested element is in position less than or equal to \( k \) is the sum of the first \( k \) probabilities, which is bounded above by \( P_1 + \cdots + P_k \).

Problem 9.6 The minimum has a simple failure rate. Let \( X_i \) be random variables with cdf \( F_i \), pdf \( f_i \) and failure rate

\[
\lambda_i(t) \equiv \frac{f_i(t)}{F_i^c(t)}, \quad t \geq 0,
\]

where \( F_i^c(t) \equiv 1 - F_i(t) \). Let \( F \) be the cdf and \( \lambda(t) \) be the failure rate of \( \min\{X_1, \ldots, X_n\} \). Then

\[
F^c(t) = \prod_{i=1}^n F_i^c(t),
\]
so that

\[ \lambda(t) = \sum_{j=1}^{n} \prod_{i \neq j} F_{i}^{c}(t) f_{j}(t) = \sum_{j=1}^{n} \lambda_{j}(t). \]

Clearly \( \lambda \) is nondecreasing whenever \( \lambda_{j} \) is nondecreasing for all \( j \).

To get a counterexample for the maximum, let \( n = 2 \) and let \( X_1 \) and \( X_2 \) be independent exponential random variables with means 1 and 2, respectively. (If they both have the same mean, the maximum is IFR.) Then

\[ F(t) \equiv P(\max\{X_1, X_2\} \leq t) = P(X_1 \leq t, X_2 \leq t) = (1-e^{-t})(1-e^{-2t}) = 1-e^{-t}-e^{-2t}+e^{-3t}, \quad t \geq 0, \]

so the density is

\[ f(t) = e^{-t} + 2e^{-2t} - 3e^{-3t}, \quad t \geq 0. \]

Thus, the failure rate is

\[ \lambda(t) = \frac{e^{-t} + 2e^{-2t} - 3e^{-3t}}{1 - [1 - e^{-t} - e^{-2t} + e^{-3t}]} = \frac{e^{-t} + 2e^{-2t} - 3e^{-3t}}{e^{-t} + e^{-2t} - e^{-3t}}. \]

Note that \( \lambda(0) = 0, \lim_{t \to \infty} \lambda(t) = 1 \) and \( \lambda(\log_{e}(3)) = 12/11 \), so that \( \lambda \) is not nondecreasing. Hence \( F \) is not IFR.

**Problem 9.9** An easy way to show the ordering for binomial random variables is to show that a Bernoulli is increasing in \( p \). Then the binomial \((n, p)\) is the sum of \( n \) IID Bernoulli random variables. Since the random variables are nonnegative, the sum is increasing in \( n \). Since each Bernoulli is increasing in \( p \), so is the sum. One can use the w.p.1 coupling here.

**Problem 9.10** Revised solution: As stated before, the key is to follow the hint. The idea is to do a special construction of the two processes on the same underlying sample space, making sure that the two processes individually have the proper finite-dimensional distributions (they individually have the proper distribution as a stochastic process). Moreover, we do the construction so that the lower process has its sample path always lying below the sample path of the upper process. As stated before, the joint distribution of the two processes is artificial. We are interested in obtaining stochastic conclusions.

However, the problem is somewhat more subtle than I was describing at first. If the upward step were always just +1, i.e., from \( i \) to \( i + 1 \) (i.e., if \( k(i) = 1 \) for all \( i \)), then a simple direct construction would be possible. Then we would be careful only when the two processes are in the same state, and otherwise let them both move as they will. However, with general \( k(i) \) it is possible that the lower process in the special construction would jump over the upper process, causing the sample path ordering to be lost.

For that reason, we stop the lower process the instant that the two processes get separated. That makes the upper process have different time than the lower process. In particular, the upper process has extra transitions, in fact, a random number of extra transitions.

The special construction is to generate the two processes in a special way whenever they are in the same state. We make the lower process have a jump down whenever the upper process has a jump down. At the same time, the upper process has a jump up whenever the lower process does. That can be achieved by using a common uniform random number \( U \), where \( U \) is uniform on \([0, 1]\), for the two processes. The upper process has \( c_i \), while the lower process has...
If \( U < c_i \), then both processes go down 1; if \( c_i < U < \bar{c}_i \), then the lower process goes down but the upper process goes up; if \( U > \bar{c}_i \), then both processes go up to \( i + k(i) \) together. When the two processes are together, we repeat the construction with a new independent uniform random variable.

The tricky step is what to do when the processes get separated. With the upward jumps defined in this problem, it is possible for the lower process to jump over the upper process. Hence, to avoid that, we stop the lower process, while we let the upper process move on. Since the upper process can go down at most 1 in any transition, the upper process cannot jump below the lower process. When they meet again, we repeat the special construction described above. But now the upper process is at some later time (larger time index) than the lower process, but always, with these altered times, the lower process has a sample path less than or equal to the sample path of the upper process.

With this construction, and these altered times, we see that the lower process will hit 0 before the upper process. Moreover, the upper process hits 0 even later, when we add in its extra times. So we can deduce that the time to hit any level from above is stochastically less for the lower process. Specifically, let \( T_{i,j}^m \) be the first passage time down to state \( j \) starting in state \( i \), assuming \( i > j \), for process \( m \), where \( m = 1 \) for the original higher process and \( m = 2 \) for the new lower process, based on \( \bar{c}_i \). Then the special construction yields

\[
T_{i,j}^2 \leq T_{i,j}^1 \quad \text{for all } i > j \text{ w.p. 1}
\]

for the special construction, which implies the associated stochastic order

\[
T_{i,j}^2 \leq_{st} T_{i,j}^1 \quad \text{for all } i > j ;
\]

i.e.,

\[
E[f(T_{i,j}^2)] \leq E[f(T_{i,j}^1)] \quad \text{for all } i > j
\]

for all nondecreasing real-valued functions \( f \) for which the expectations are well defined. (Even though we did not do so, it is natural to use \( \tilde{T}_{i,j}^m \) for the random variables involved with the special construction, and other variables \( T_{i,j}^m \) for the random variables without the special construction. All three relations above apply to the \( \tilde{T}_{i,j}^m \) variables, but the last two also apply to the \( T_{i,j}^m \) variables.)

Actually Ross asked about an easier quantity, the \( f_i \), the probability of ever hitting 0 starting in state \( i \). Clearly this probability is greater for the lower process. Because, if the upper process ever hits 0, then so necessarily does the lower process, even sooner. Our comparison above involving the times to hit any lower level is a stronger comparison than made by Ross.

**Problem 9.11** Recall that

\[
N(\mu, \sigma^2) \overset{d}{=} \mu + \sigma N(0, 1).
\]

Clearly,

\[
\mu_1 + \sigma N(0, 1) \leq \mu_2 + \sigma N(0, 1)
\]

when \( \mu_1 \leq \mu_2 \).

There is no stochastic order as \( \sigma^2 \) increases. If we just change \( \sigma^2 \), then we have convex stochastic order; see Section 9.5.

**Problem 9.13** We generate the renewal processes in a special construction so that the points in the second process are a subset of the points in the first process. We generate points
in process 1 in any manner. Suppose that a point occurs in the first process at some time, which is $t$ time units after the last point in process 1. At this time, we observe the elapsed time since the last point in the second process. Suppose that elapsed time is $s$. Then we let a point also occur in the second process at this time with probability $\lambda_2(s)/\lambda_1(t)$. Otherwise, there is no point in the second process, and the age increases. It can be shown that this construction gives process two the correct distribution. Clearly, the special construction makes the points in process 2 a subset of the points in process 1.

**Problem 9.18** The key is to write the ccdf in terms of the hazard-rate (or failure-rate) function:

$$P(X > t) = e^{-\int_0^t \lambda(x)\,dx}.$$  

Note that

$$\frac{P(X_1 > t)}{P(X_1 > s)} = \frac{e^{-\int_0^t \lambda_1(x)\,dx}}{e^{-\int_0^s \lambda_1(x)\,dx}} = e^{-\int_s^t \lambda(x)\,dx}.$$  

Then the result follows, because

$$\lambda_1(x) \geq \lambda_2(x) \quad \text{for all } x \quad \text{if and only if} \quad \int_s^t \lambda_1(x)\,dx \geq \int_s^t \lambda_2(x)\,dx \quad \text{for all } t > s.$$  

Extra Problem 1. In the *Resource Sharing* paper it states that monotone likelihood ratio (MLR) order implies stochastic order for random variables taking nonnegative-integer values. Prove it.

Let $p_k = P(X = k)$ and $q_k = P(Y = k)$ for $k \geq 0$. The MLR order $X \leq_r Y$ can be expressed as

$$\frac{p_k}{p_{k-1}} \leq \frac{q_k}{q_{k-1}} \quad \text{for all } k.$$  

First note that $X \leq_r Y$ implies that $p_0 \geq q_0$. To see that, observe that

$$p_k = \frac{p_k}{p_{k-1}} \frac{p_{k-1}}{p_{k-2}} \cdots \frac{p_2}{p_1} \frac{p_1}{p_0}$$

and

$$q_k = \frac{q_k}{q_{k-1}} \frac{q_{k-1}}{q_{k-2}} \cdots \frac{q_2}{q_1} \frac{q_1}{q_0}.$$  

Thus

$$1 = \sum_k p_k = p_0 \sum_k c_k \quad \text{and} \quad 1 = \sum_k q_k = q_0 \sum_k d_k,$$

where $c_k \leq d_k$ for all $k$. Since

$$\sum_k c_k \leq \sum_k d_k,$$

we must have

$$p_0 \geq q_0.$$  

Then the MLR ordering implies that there is a $k^*$ such that

$$p_k \geq q_k \quad \text{for all } k < k^*,$$

while

$$p_k \leq q_k \quad \text{for all } k \geq k^*.$$  

5
As a consequence,
\[ \sum_{i=0}^{k} p_i \geq \sum_{i=0}^{k} q_i \quad \text{for all} \quad k < k^* , \]
while
\[ \sum_{i=k}^{\infty} p_i \leq \sum_{i=k}^{\infty} q_i \quad \text{for all} \quad k \geq k^* . \]

Combining the last two inequalities, we obtain
\[ X \leq_{st} Y . \]

Extra Problem 2. Prove or disprove: For random variables taking values in the nonnegative integers, MLR order is equivalent to stochastic order.

This is false; MLR order is stronger than stochastic order. Let
\[ p_0 \equiv P(X = 0) = 0.2, \quad p_1 \equiv P(X = 1) = 0.1 \quad \text{and} \quad p_2 \equiv P(X = 2) = 0.7 \]
and
\[ q_0 \equiv P(Y = 0) = 0.3, \quad q_1 \equiv P(Y = 1) = 0.4 \quad \text{and} \quad q_2 \equiv P(Y = 2) = 0.3 . \]
Then \( X \geq_{st} Y \), but we do not have MLR order:
\[ \frac{p_0}{q_0} > \frac{p_1}{q_1} , \]
while
\[ \frac{p_1}{q_1} < \frac{p_2}{q_2} . \]