SOLUTION 6 (18.05, SPRING 2003)

H6-1 a) By the symmetry among the balls, the probability that each particular ball is red is $\frac{6}{24+6} = \frac{1}{5}$. Now, by linearity of expectations the expected number of red balls each of the players has is $E_i = 6 \cdot \frac{1}{5} = \frac{6}{5}$. Again, by linearity of expectations the first two players have the expected number of $E = E_1 + E_2 = \frac{6}{5} + \frac{6}{5} = \frac{12}{5}$ red balls.

b) By the linearity of expectations, the expected number of the difference in the number of red balls equals the difference in the expectations: $E' = E_1 - E_2 = 0$.

One can prove this also by noting the symmetry between two players (the expected number of red balls is independent on the order in which they collect them from the urn).

H6-2 Fix the order in which people are standing in line. The probability P that the first person knows the second, the second knows the first, etc. is

$$p \cdot p \cdot p \cdots p \ (n\text{times}) \ = \ p^{n-1}.$$

There are $n!$ ways to order these people. By linearity of expectations, we obtain $E = n! \cdot p^{n-1}$ is the expected number of ways in which they can form a line as in the problem.

Say, the party has $n = 20$ people and random two know each other with probability $p = 0.1$. We have:

$$E = n! \cdot p^{n-1} \approx 0.243$$

so this line is relatively unlikely to exist. If, however, $p = 0.2$, then

$$E = n! \cdot p^{n-1} \approx 127554.13,$$

i.e. this line is *very* likely to exist.

H6-3 Say, there are N days in a year, and assume that birthdays are uniformly distributed among these N days. The probability that 3 particular people have a birthday on a given day is $(1/N)^3$. Summing this probability over all days, we obtain $p = 1/N^2$ probability that 3 particular people have a birthday on the *same* day. There are $\binom{n}{3}$ ways to choose 3 people out of n people. Now, by linearity of expectations the expected number of triples of people who share the same birthday is

$$E = \binom{n}{3} \cdot \frac{1}{N^2}$$

For a class of $n = 60$ people, this expectation is

$$E = \binom{60}{3} \cdot \frac{1}{365^2} = \frac{34220}{133225} \approx 0.26$$

Trying various n one finds that for $n = 93$ this expectation is ≈ 0.97, but for $n = 94$ we have $E \approx 1.01$.

H6-4 Do these facts contradict each other - No! One can come up with any such income distribution as long the percentage of people making $\geq x$ amount is *decreasing* as x is increasing (for the finitely many claims as in the problem).

Now, suppose the average income is $100K$. The first claim is definitely false since if true it would imply that the average income is at least

$$0.40 \cdot 300K = 120K > 100K$$

Same for the second claim:

$$0.15 \cdot 900K = 135K > 100K$$

but not for the third claim:

$$0.01 \cdot 3000K = 30K < 100K$$

Thus in fact the third statistician's claim may as well be true. Defending the statistician in the first two claims is easy: you argue that he made 20% and 35% error in these cases, respectively. You argue in the first case that $2/3 \approx 67\%$ of people earn nothing ($0K$) and the remaining 33% earn exactly $300K$. In the second case you argue that $1/9 \approx 89\%$ of people earn nothing ($0K$) and the remaining 11% earn exactly $900K$. Thus, you argue it's a small error in statistician's sampling work.

Now, each of these arguments (viewed independently) is implausible, but there is not enough information to reject these claims. On the other hand, together these arguments contradict each other. Basically, this statistician should be fired. :)

H6-5 First, if the average temperature in the first week of March is x degrees, the next week it's $x + 5$, then $x + 10$, and finally $x + 15$. The average of these is

$$\frac{x + (x + 5) + (x + 10) + (x + 15)}{4} = x + 7.5 \; = \; 50$$

Thus $x = 42.5$. Now, by Markov inequality, taking $\lambda = 80/42.5$ we get

$$P(\text{the temperature} \geq 80) \; = \; P(\text{the temperature} \geq \lambda \cdot x) \leq \frac{1}{\lambda} \approx 0.53$$

In other words, the probability that the temperature in the first week of March is ≥ 80 is *at most* 53%. (I wish...)