Statistics 7 Practice Midterm Solution

1. (a) Right skewed.
   (b) About 55.
   (c) About 45, so lower than that in part (b), the last five points are all above the line.

2. (a) Divide a 20-day period into 10 two-day pairs. Randomly assign the 20 subjects to the 20 days. Within each two-day pair flip a coin and randomly assign one day (and its subject) to the 70 degree treatment and the other day in the pair to the 90 degree treatment.
   (b) Again divide a 20-day period into 10 two-day pairs and randomly assign *one* subject to each two-day pair. Each subject will preform the task in both days in their assigned two-day pair, one day at 70 degrees and the other at 90 degrees. For each subject flip a coin to decide if the first or the second day will be the 70 degree treatment.
   (c) The matched pairs design is better because its effective control of variability among the subjects makes the treatment vs. control comparison more precise.
   (d) Yes. If all 70 degree subjects are run on the same day and all 90 degree subjects in the another day, the treatment effect of temperature is confounded with a possible day effect. (For example, something else may go wrong on one of the days to slow down the work.)

3. (a) Father's and son's height exhibit an approximate linear relationship. As father's heights increase, the conditional mean of son's heights given father's heights increase linearly, but the conditional variance is roughly the same. Note the conditional distribution of son's heights appears to be normal.
   (b) Son's Height=36.1172 + .4854401*Father's Height. Except for very short and very tall fathers, this line is an adequate description of the relationship between the two variables. 1 inch increase in father's height is associated with a .4854401 inch increase in son's height, on average.
   (c) $r * s_y/s_x = .4854401$, $r = .4854401 * 2.671891/2.609882 = 0.4969$, and R^2, 0.2469, is the amount of variation explained by father's heights.
   (d) $36.1172 + .4854401 * 72 = 71.07$ inch
   (e) Shorter. This is an example of the regression effect.
   (f) $\sqrt{(1 - R^2) * s_y^2} = 2.27$ inches

4. (a) $X \sim \text{binomial}(240, 0.1)$, the number of successes in 240 independent trials, with constant success probability of 0.1.

   (b) X: $\mu_X = 240 * 0.1 = 24$, $\sigma_X^2 = 240 * 0.1 * 0.9 = 21.6$. $\hat{p} = X/240$: $\mu_{\hat{p}} = 0.1$, $\sigma_{\hat{p}}^2 = 0.1 * 0.9/240 = 0.000375$.

   (c) $P(X \geq 30) = P(Z \geq \frac{30-24}{\sqrt{21.6}}) = P(Z \geq 1.20) = 0.098$, where Z is a standard normal random variable. Assumption: X is approximately normally distributed (or the central limit theorem holds). Justification: $np > 10$.

5. (a) and

   (b) :

Ann	Bob	X
A	A	0
A	B	2
A	C	-3
A	D	0
B	A	-2
B	B	0
B	C	0
B	D	3
C	A	3
C	B	0
C	C	0
C	D	-4
D	A	0
D	B	-3
D	C	4
D	D	0

   (c) :

X	Probability
-4	1/16
-3	2/16
-2	1/16
0	8/16
2	1/16
3	2/16
4	1/16

   (d) $(1 + 2 + 1)/16 = 1/4$

   (e) The probability that no money changes hands in each of the four rounds is $(1/2)^4 = 1/16$. So the chance that money changes hands in at least one found is 15/16.