STAtistics / Mr. Hansen Name: KEY
12/9/2008 Battery bonus (Mr. Hansen's use only): ________

Cumulative Test through Chapter 8 (100 points)

- Illegible or cryptic explanations may not earn full credit, even if your answers are numerically correct. In particular, the letter z must always be crossed.
- Decimal approximations should always include at least 3 significant digits.
- Do not mark on your formula sheets. These are to be reused all year.
- If you happen to have spare batteries, leave them in plain view for a 1-point bonus.

Part I: Always, Sometimes, Never (2 pts. each)
In the small blank, write A if the statement is always true, S if sometimes true, or N if never true.

right ($n=11$)

8 S 1. If A and B are events, then $P(A \cap B) = P(A) \cdot P(B)$.

5 A 2. If A and B are events, then $P(A \cap B) = P(A) + P(B) - P(A \cup B)$.

8 A 3. Residuals for a LSRL have a sum of 0.

8 N 4. If $s_x \neq 0$, there is more than one line that minimizes $\sum (y - \hat{y})^2$.

9 N 5. "House effect" refers to sampling error, i.e., random variations between samples.

8 A 6. An r value close to –1 implies strong linear correlation.

3 S 7. A statistically significant result occurs by chance. (When this happens, we call it "Type I error.")

5 S 8. A P-value of less than 0.05 constitutes statistical significance.

3 N 9. The P-value tells us the probability that chance alone caused a certain result to appear. (Nobody can tell us this probability.)

6 A 10. The P-value is a conditional probability, namely the probability of finding results as extreme as or more extreme than the ones observed, given that chance alone is the only force at work and there is no actual difference between the experimental and control groups for the parameter being sampled.

Part II: Definitions (4 pts. each).
Give a concise, reasonably accurate definition of each term.

11. statistical significance
magnitude of a difference considered too large to be plausibly explained by chance

12. ordinal scale
a variable or measurement instrument whose possible values may or may not look like numbers, but for which an inherent order exists

KEY p.2

13. normal quantile plot

a scatterplot of x values plotted against their percentiles' z-scores: straight line denotes normality, while a bend to the R or L denotes R or L skewness, respectively

14. independent (adj.)

meaning for events, "satisfying $P(A \cap B) = P(A) \cdot P(B)$";

for random variables, "satisfying a condition whereby knowledge of the value (or value interval) of one does not affect the conditional probability distribution of the other"

15. random variable

a numeric outcome of a random process; or, a numeric variable X whose possible values x_i satisfy either $\sum p_i = 1$ (discrete) or $\int_{-\infty}^{\infty} f(x)\,dx = 1$ (cont.) for nonneg. probabilities p_i or density fcn. f, where $p_i \ngtr 1$.

Part III: Probability (60 pts., 6 pts. per numbered question): JUSTIFY ANSWERS! Proper notation is required for full credit. Answers alone are worth 2 points per question. The remainder of the points are for proper notation and justification. In general, you must furnish setup (often a formula), plug-ins, and answer.

For questions 16-21, assume that Carl sinks approximately 82% of his free throws and that the shots are independent.

16. Define a suitable random variable for the number of tries needed for Carl to sink his first free throw. Identify the setting by name, and explain how you know that all the conditions are met.

Let X = # of tries needed for Carl to sink first free throw

X is geometric with parameter $p = .82$

- Indep. trials (given)
- p is fixed (given)
- Vbl. of interest is # of attempts needed for first success
- Only two possibilities per trial (success or failure)
- n is undefined

17. Compute the probability that Carl sinks his first free throw in fewer than 4 tries.

$$P(X < 4) = P(X \leq 3) = p + pq + pq^2 = .994$$

18. Define a suitable random variable for the number of successful free throws Carl obtains in 18 trials. Identify the setting by name, and explain how you know that all the conditions are met.

Let Y = # of successes in 18 shots

Y is $B(18, .82)$ since

- indep. trials (given)
- p is fixed (given)
- vbl. of interest is # of successes in 18 tries
- only 2 possibilities per trial (success or failure)
- $n = 18$ is fixed

KEY p.3

19. Compute the probability of more than 12 but fewer than 15 successes in 18 tries.

Let $p = .82$, $q = .18$

$$P(12 < Y < 15) = P(Y=13) + P(Y=14)$$
$$= \binom{18}{13} p^{13} q^{5} + \binom{18}{14} p^{14} q^{4}$$
$$= .12269 + .1996$$
$$= .3223$$

20. Compute the expected number and variance for the number of successes in 18 tries.

$$\mu_Y = E(Y) = np = 18(.82) = 14.76$$

$$Var(Y) = \sigma_Y^2 = npq = 18(.82)(.18) = 2.6568$$

21. Compute the expected number of shots needed in order for Carl to obtain his first success.

$$E(X) = \mu_X = \frac{1}{p} = \frac{1}{.82} = 1.2195$$

22. The incidence of TB in a large city is 1.2%. An SRS of 700 residents is selected. Explain why the binomial distribution for the number of TB-afflicted people in the sample is not correct but is nevertheless close enough for estimation purposes.

SRS means "without replacement," hence not independent. However, in a large city, the difference between replacing or not between trials is negligible.

$$P(\text{at least one with TB}) = 1 - P(\text{none with TB})$$
$$= 1 - (1 - .012)^{65}$$
$$= .54375$$

23. Under the same givens as in #22, compute the probability that an SRS of 65 residents contains at least one TB-afflicted person.

24. The probability of "two pair" on the deal in 5-card draw poker is 123552/2598960, as we computed in class some weeks ago. Prove this result. ("Two pair" is defined as two cards of one value, two cards of another value, and one "junk card" drawn from the remaining 44. There are 52 cards in a standard deck, consisting of 13 values and 4 distinguishable cards of each value.)

choose 2 values; choose 2 of one value; choose 2 of other value; choose junk card

$$\frac{\binom{13}{2}\binom{4}{2}\binom{4}{2}\binom{44}{1}}{\binom{52}{5}} = \frac{123552}{2598960} \quad \text{as claimed}$$

total # of 5-card poker hands

25. A woman plays 600 hands of 5-card draw poker. Compute the probability that the number of times she receives "two pair" on the deal is 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30. If possible, please use a time-efficient method.

Let X = # of "two pair" hands in 600 trials.

Note: $X = B\left(600, \frac{123552}{2598960}\right)$

$$P(21 \le X \le 30) = P(X \le 30) - P(X \le 20)$$

$$= .656845 - .056284$$

$$= .60056$$