Statistics for the Life Sciences

Math 20340 Section 01, Fall 2008

Homework 5 Solutions

- **6.32**: Say the customer asks for tension X. If the stringer sets his machine to have average tension X, then half the time the tension will be too low. If he sets it lower than X, then on average it will be less than X, and overall *more* than half the time. So to be sure that only 5% of the time he gets a racket off the machine that has string tension below X, it seems clear that he needs to set the tension average to be (quite a bit) higher than X.

  Let μ be the tension the stringer chooses as the average. Then the tensions x of rackets off the machine are normally distributed with mean μ and standard deviation 2. We want to make sure that $P(x \leq X) = .05$. Converting to standard normal, this is the same as $P(z \leq (X - \mu)/2) = .05$. From a standard normal table, we find that $P(z \leq -1.645) = .05$. So we need to choose μ so that $(X - \mu)/2 = -1.645$ or $X - \mu = -3.29$. In other words, the stringer must set the average tension to be 3.29 lbs (or 1.645 standard deviations) higher than the requested tension; anything lower, and more than 5% of the rackets will be below the customers specification.

- **6.33**: Let x be amount shopper has spent. x is normal with mean 85, std dev 20, so $(x - 85)/20$ is a standard normal.

  - $P(x > 95) = P(z > (95 - 85)/20) = P(z > .5) = 1 - P(z \leq .5) = 1 - .6915 = .3085$
  - $P(95 \leq x \leq 115) = P(x \leq 115) - P(x \leq 95) = P(z \leq 1.5) - P(z \leq .5) = .9332 - .6915 = .2417$
  - For a single shopper: $P(x > 115) = P(z > 1.5) = 1 - P(z \leq 1.5) = 1 - .9332 = .0668$. If the two shoppers are independent, the probability of both spending more than \$115 is $(.0668)^2 = .0044$...

- **7.19**: In each case mean of $\bar{x}$ is population mean, and standard deviation of $\bar{x}$ equals population std dev over square root of sample size.

  - **a**: Mean 10, std dev $.5$
  - **b**: Mean 5, std dev $.2$
  - **c**: Mean 120, std dev $1/\sqrt{8} = .35355$...

- **7.20**: See box of page 266, "How do I decide when the sample size is large enough"?

- **a**: $\bar{x}$ will be *exactly* normal in all cases
- **b**: $\bar{x}$ will be approximately normal for a) and b); but we can't say anything about c), since n is too small

- **7.25**:
  - **a**: Mean 106, std dev 2.4
  - **b**: $P(\bar{x} > 110) = P((\bar{x} - 106)/2.4 > (110 - 106)/2.4) = P(z > 1.667) = 1 - P(z \leq 1.667) = 1 - .9525 = .0475$
  - **c**: $P(102 \leq \bar{x} \leq 110) = P(-1.67 \leq z \leq 1.67) = .9525 - .0475 = .905$

- **7.27**:
  - **a**: Air temperature, time at which measurement is taken, variations in amounts of the various substances introduced initially, ...
  - **b**: Large number; variability of average error (measured by std dev) = variability of single measurement over square root of number of measurements; larger number of measurements leads to smaller variability.

- **7.29**: Think of a cubic foot of water as being made up of many (1728) cubic inches. Total number of bacteria is equal to sum of numbers in each square inch. It's reasonable to assume that the numbers in different square inches are independent, so total is sum of results of large number of independent experiments, each with same mean and std dev. Central limit theorem applies to say that the sum is approximately normal.

- **7.31**: Let x_i be amount of Potassium in banana i. Each of x_1, x_2 and x_3 are independent normal random variables with mean 630 and std dev 40.
  - **a**: $T = x_1 + x_2 + x_3$; by our sum variant of central limit theorem this has mean $3 * 630 = 1890$ and std dev $\sqrt{3} * 40 = 69.28$.
  - **b**: $P(T > 2000) = P(z > 1.59) = .0559$ (2000 is 1.59 standard deviations above 1890).

- **7.33**:
  - **a**: Distribution of sample mean for sample of size 130 has mean 98.6, std dev $.8/\sqrt{130} = .07...$. 98.25 is .35 below mean, or 5 standard deviations, so probability is very low (very close to zero).
  - **b**: Very unlikely.

- **7.37**:
  - **a**: mean .3, std dev .0458...
  - **b**: mean .1, std dev .015...

- **c**: mean .6, std dev .03...

- **7.38**:
  - **a**: No, $np = 2.5 \leq 5$
  - **b**: Yes, $np = 7.5$, $nq = 67.5$, both > 5
  - **c**: No, $nq = 2.5 \leq 5$

- **7.41**: $SE(\hat{p}) = \sqrt{\hat{p}\hat{q}/n}$ in all cases.
  - **a**: .0099...
  - **b**: .03
  - **c**: .0458...
  - **d**: .05
  - **e**: .0458...
  - **f**: .03
  - **g**: .0099...
  - **h**: Max at $p = .5$; close to zero for p close to 0 or 1

- **7.47**:
  - **a**: $p = .75$, $n = 200$, so mean of $\hat{p}$ is $p = .75$ and standard deviation is $\sqrt{pq/n} =$ .03...; distribution is approximately normal
  - **b**: Greater than 80% is .05 above mean, or 1.67 standard deviations. Probability of this is .0475.
  - **c**: 95% of the time the sample proportion will be within ± 1.96 std devs of mean, so in range $.75 - 1.96 * .03$ to $.75 + 1.96 * .03$ or .6912 to .8088.