Homework #7 Solutions

5.3, 5.7, 5.12, 5.13, 5.31, 5.35

Problem 5.3

(a) The number of observations (50) is fixed. It is also reasonable to assume that each students exam result is independent and the probability of passing is the same for each student. We are interested in a R.V $X \equiv$ "number of students who pass out of 50". The binomial setting is reasonable for this senario

(b) Although the number of observations is fixed (10) and we are interested in a r.v. $X \equiv$ "number of questions out of 10 correct", it is not binomial because the student gets extra help between questions when they make a mistake. Therefore the observations are not independent.

(c) Again the number of observations is fixed (10) and we are interested in a random variable $X \equiv$ "number of times substance disolves in 10 experiments". However, it does not seem sensible to assume that the probability of melting remains the same each time because the temperature is being changed. Therefore the binomial is not appropriate

Problem 5.7

from 5.5 we learn that

X ≡ "number of Hispanics on committee of 15 members"

is Binomial with $n = 15$ $p = 0.3$

a) for a Binomial r.v the mean is $\mu_X = np$

so
$$\mu_X = 15(.3) = 4.5$$

b) for a Binomial r.v. the standard deviation is

$$\sigma_X = \sqrt{np(1-p)}$$ so

$$\sigma_X = \sqrt{15(.3)(1-.3)} = 1.7748 \text{ (4dp)}$$

c) now $p = 0.1$ so

$$\sigma_X = \sqrt{15(.1)(.9)} = 1.1619 \text{ (4dp)}$$

with $p = .01$

$$\sigma_X = \sqrt{15(.01)(.99)} = 0.3854 \text{ (4dp)}$$

we learn that the standard deviation gets smaller as p approaches 0

Problem 5.12

(a) We have a fixed sample size 200. A respondent can either answer that they "are committed to eating nutritious food when eating away from home" or that they "aren't committed". Since a SRS is being used it seems reasonable to assume that responses are independent and p remains same between respondents. So the binomial setting seems reasonable. If we assume the national result then $p = .4$ so Binomial(200, .4) is reasonable distribution for

$X \equiv$ "number of respondents out of 200 who 'are committed' "

(b) We want $P(75 \leq X \leq 85)$

because $np = 200(.4) = 80 > 10$
and $n(1-p) = 200(.6) = 120 > 10$

④

it is reasonable to use a normal distribution approximation to the binomial distribution. In particular,

X is approximately normal with mean $\mu = np = 200(.4) = 80$

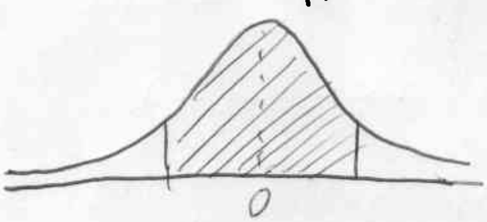
and standard deviation $\sigma = \sqrt{np(1-p)} = \sqrt{200(.4)(.6)} = \sqrt{48} = 6.9282$

So

$$P(75 \le X \le 85) \approx P\left(\frac{75-80}{6.9282} \le \frac{X-80}{6.9282} \le \frac{85-80}{6.9282}\right)$$

because of the approximation

$$= P(-0.72 \le Z \le 0.72)$$



$$= P(Z \le .72) - P(Z \le -.72)$$

$$= .7642 - .2358$$

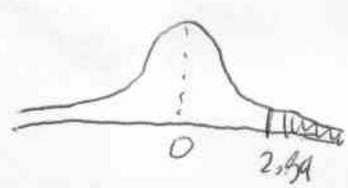
$$= .5284$$

c) What we want to know is

$$P(X \ge 100) \approx P\left(\frac{X-80}{6.9282} \ge \frac{100-80}{6.9282}\right)$$

again due to the approximation

$$= P(Z \ge 2.89)$$



$$= 1 - P(Z < 2.89)$$

$$= 1 - .9981$$

$$= .0019$$

since this is small it is not unreasonable to think p might be lower.

Problem 5.13

(a) $\hat{p} = \dfrac{X}{n} = \dfrac{62}{100} = .62$

(b) Because $np = 100(.67) = 67 > 10$
and $n(1-p) = 100(.33) = 33 > 10$

it is reasonable to use a normal approximation.
In particular $\hat{p}$ has (approximately) normal distribution with mean $\mu_{\hat{p}} = p = .67$ and standard deviation $\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}} = \sqrt{\dfrac{.67(1-.67)}{100}}$

$= .0470$ (4dp)

So

$$P(\hat{p} \le 0.62) \approx P\left(\frac{\hat{p} - .67}{.0470} \le \frac{.62 - .67}{.0470}\right)$$

because we use normal approximation

$$= P(Z \le -1.06)$$

$$= .1446$$

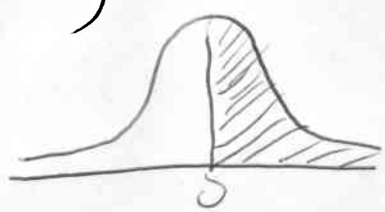
(c) Since the probability of getting 62 or smaller (ie $\hat{p} = .62$ or smaller) just by chance if $p = .67$ is not that small the survey does not support the conclusion that p is smaller at Harvard College.

Problem 5.31

(a) we are told that $X \equiv$ "a 12 graders test score" is approximately normal with mean $\mu = 300$ and standard deviation $\sigma = 35$. so

$$P(X > 300) = P\left(\frac{X-300}{35} > \frac{300-300}{35}\right)$$

$$= P(Z > 0)$$



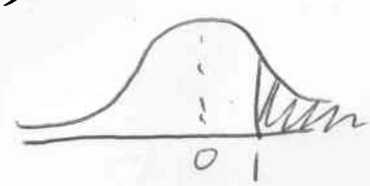
$$= 1 - P(Z < 0)$$

$$= 1 - .5$$

$$= .5$$

$$P(X > 335) = P\left(\frac{X-300}{35} > \frac{335-300}{35}\right)$$

$$= P(Z > 1)$$



$$= 1 - P(Z < 1)$$

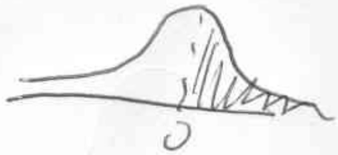
$$= 1 - .8413$$

$$= .1587$$

(b) Now we are interested in the mean of four 12 graders. Since each individual measurement is normal (approximately) so is the sample mean. Note that

$$\mu_{\bar{x}} = \mu = 300 \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{35}{\sqrt{4}} = \frac{35}{2} = 17.5$$

$$P(\bar{x} > 300) = P\left(\frac{\bar{x}-300}{17.5} > \frac{300-300}{17.5}\right)$$
$$= P(z > 0)$$



$$= 1 - P(z < 0)$$
$$= 1 - .5$$
$$= .5$$

$$P(\bar{x} > 335) = P\left(\frac{\bar{x}-300}{17.5} > \frac{335-300}{17.5}\right)$$
$$= P(z > 2)$$

$$= 1 - P(z < 2)$$
$$= 1 - .9772$$
$$= .0228$$

<u>Problem 5.35</u>

we are told that $x \equiv$ "blood glucose level for Shelia" is Normal with $\mu = 125$ $\sigma = 10$

to be diagnosed as diabetic we need $X > 140$. so

$$P(X > 140) = P\left(\frac{X-125}{10} > \frac{140-125}{10}\right)$$
$$= P(Z > 1.5)$$

$$= 1 - P(Z < 1.5)$$
$$= 1 - .9332$$
$$= .0668$$

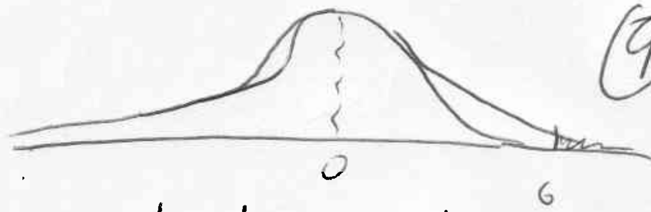
b) Now we are interested in the mean of 4 different measurements. The sample mean $\bar{X}$ will also have normal distribution, but its mean is

$$\mu_{\bar{X}} = \mu = 125$$

and standard deviation is

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{4}} = 2.5, \text{ so}$$

$$P(\bar{X} > 140) = P\left(\frac{\bar{X}-125}{2.5} > \frac{140-125}{2.5}\right)$$



$= P(Z > 6)$

$= 1 - P(Z < 6)$

$< 1 - P(Z < 3.9) = 1 - .9998$ ← largest value in table

note change of sign to "less than" rather than "equals"

$= .0002$

ie it is much less likely she will be diagnosed as diabetic if we use the mean of 4 observations rather than a single observation.