STA286 exam, April 2018: Notes and partial solutions

General notes

Students were told that the exam covered the following textbook sections:

- 5.1-5.5, 6.1-6.7, 6.10, 7.3, 8.1-8.8, 9.1-9.6, 9.8, 9.14, 10.1-10.6, 10.10, 10.11, 11.1-11.6. (Excl. § 11.12.)

Professor Kundur supplied questions 1 to 5, Professor Ebden questions 6 to 18.

Some of the advice given to graders

- For full marks, students need to arrive at the right answer and to show their work (unless the question says otherwise)
- If the student's answer is wrong, please look through it until you spot the mistake(s). This will help you decide on part-marks. You don't need to write comments about their mistakes

Partial solutions

1. (See Walpole 5.31.)

2. (See Walpole 5.70.)

3. (a) (See Walpole 5.1.) (b) (8 marks)

4. (See Walpole 6.3.)

5. (a) (Four marks) (b) (See Walpole 6.45.)

6. $M'_X(t) = k\, e^t\, e^{k(e^t-1)}$. The mean, $\mathrm{E}(X)$, is given by $M'_X(0) = k(1)e^0 = k$.

Graders beware for shortcuts that bypass the mgf.

7. See Walpole 8.2. Answers: 9, 6, and 5.

8. $\mathrm{var}(X_i) = \mathrm{E}[(X_i - 0)^2] = \int_{-1/2}^{1/2} x^2\, dx = \frac{x^3}{3} \,\Big|_{-1/2}^{1/2} = 1/12$

By CLT, $\lim_{n\to\infty} P\left(\frac{S_n}{\sigma\sqrt{n}} \le x\right) = \Phi(x)$. So,

$$P\left(\frac{S_n}{\sqrt{48}/\sqrt{12}} \le x\right) = P\left(S_n \le 2x\right) = \Phi(x)$$

$$\begin{aligned} P\left(S_n \le 3\right) &= \Phi(1.5) \\ P\left(S_n \le 3\right) &\approx 0.9332 \quad \text{from the aid sheet} \\ P\left(S_n > 3\right) &\approx 0.0668 \end{aligned}$$

9. D, normal and χ^2

10. See Walpole 9.3. We have $\bar{x} = 310$, $\sigma = 1.5$, $n = 75$, and $z_{0.05} = 1.645$. The CI is

$$310 \pm (1.645)1.5/\sqrt{75} \quad = \quad 310 \pm 0.285 \quad = \quad (3.0972, 310.28)$$

Because the question only asks for a minimum of 2 significant figures, you may also award full marks for the unusual answer of (310, 310).

11. See Walpole 9.39. We have $\bar{x}_1 = 84$ & $\bar{x}_2 = 77$, $\bar{s}_1 = 4$ & $\bar{s}_2 = 6$, and $\bar{n}_1 = 12$ & $\bar{n}_2 = 18$. So

$$s_p = \sqrt{\frac{(11)4^2 + (17)6^2}{28}} \approx 5.305 \qquad \text{and} \qquad t_{0.025} = 2.048 \quad \text{with 28 degrees of freedom}$$

So the CI is

$$(84 - 77) \pm (2.048)5.305\sqrt{1/12 + 1/18} \quad = \quad 7 \pm 4.049 \quad = \quad (2.951, 11.049)$$

Because the question only asks for a minimum of 2 significant figures, you may also award full marks for (3.0, 11).

12. See Walpole 9.81. Namely,

$$\begin{aligned} L(x;p) &= f(x;p) = p^x(1-p)^{1-x} \\ \log L &= x \log p + (1-x)\log(1-p) \\ \frac{\partial \log L}{\partial p} &= \frac{x}{p} + \frac{(1-x)(-1)}{1-p} \\ 0 &= \frac{x}{\hat{p}} - \frac{1-x}{1-\hat{p}} \\ 0 &= x - x\hat{p} - \hat{p} + x\hat{p} \\ \hat{p} &= x \end{aligned}$$

Therefore, given an observed $x = 1$, the maximum likelihood estimator is x. Also accept: "the maximum likelihood *estimate* is 1."

13. See Walpole 10.1. She could conclude that less than 10% of the public are allergic to the product when, in fact, 10% or more are allergic.

14. See Walpole 10.29. The hypotheses are:

$$H_0 : \mu = 40 \qquad \text{and} \qquad H_1 : \mu \neq 40$$

We have $n = 20$, $\bar{x} = 38.1$, $s = 4.3$, $t_{0.025} = 2.093$ with 19 degrees of freedom, and

$$t_{\text{obs}} = \frac{38.1 - 40}{4.3/\sqrt{20}} \approx -1.98$$

Decision: Do not reject H_0.

Also accept: "We are unable to conclude that the average test time differs from 40 minutes," or a confidence-interval approach (as per the strong wording on p. 339 of the textbook) because the question didn't specifically disallow this.

15. See Walpole 10.80. The hypotheses are:

H_0 : The grade distribution is uniform and H_1 : The grade distribution isn't uniform

We have $\chi^2_{\text{crit}} = 9.488$ with 4 degrees of freedom, and

$$\begin{aligned}
\chi^2 &= \sum_{i=1}^{5} \frac{(o_i - e_i)^2}{e_i} \quad \text{where } e_i = 200/5 = 40 \\
&= \frac{1}{40}\left[(28-40)^2 + (36-40)^2 + 24^2 + 0^2 + 8^2\right] \\
&= 800/40 \\
&= 20
\end{aligned}$$

Decision: Reject H_0.

Also accept: "We conclude that the distribution of grades isn't uniform."

16. D, squared vertical distances.

17. B, error.

18. See Walpole 11.6 and 11.23(a).

(a) Using $b_1 = \frac{S_{xy}}{S_{xx}}$ and $b_0 = \bar{y} - b_1\bar{x}$, where $S_{xx} = 8$ etc, we get $\hat{y} = b_0 + b_1 x = 59.75 - 1.25x$

(b) $s^2 = \text{SSE}/(3-1) = \sum_{i=1}^{3}(y_i - \hat{y}_i)^2 = (29-28.5)^2 + (25-26)^2 + (24-23.5)^2 = 1.5$

(c) $\hat{y} = b_0 + b_1(30) = 22.25$

(d) A confidence interval on Y at a value x_0 is:

$$(b_0 + b_1 x_0) \pm t_{\alpha/2}\, s\sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}$$

where the t-distribution has $n-2$ degrees of freedom. So,

$$22.25 \pm 6.314\sqrt{1.5}\sqrt{\frac{1}{3} + \frac{9}{8}} \quad \approx \quad 22.25 \pm 9.34 \quad \approx \quad (12.9,\ 31.6)$$

This is the correct answer. Subtract 2 marks if they calculate successfully a *prediction* interval:

$$(b_0 + b_1 x_0) \pm t_{\alpha/2}\, s\sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}$$

where again the t-distribution has $n-2$ degrees of freedom. This leads to $22.25 \pm 12.12 \approx (10.1,\ 34.4)$.