STAT/CS 94 Fall 2015 Adhikari Midterm Exam Solutions

This solutions sheet has short answers to the midterm questions, as we would have answered them. The Gradescope rubric has more detailed information about the grading of each problem.

Problem 1

Answer:
(a) `len(cards)`

(b)
```
def last_four(n):
    return "Card ending in " + str(n)[-4:]
```

Problem 2

Answer: A: 2; B: 4; C: 5

A and C: The empirical histogram of a large random sample looks like the population; bigger samples look more like the population than smaller ones do.

B: The minimum of 100 random draws will be pretty small, certainly not balancing to the right of 50 as the rest of the histograms are.

Problem 3

Answer:
(a) mean: 60.07; median: 56.11; SD: 26.71

The histogram won't balance at 26 or below. Because of the right hand tail, the mean will be bigger than the median, so the mean is 60.07 and the median is 56.11. Chebychev says that "mean plus or minus 2 SDs" includes at least 75% of the data, which rules out 5.01 as a candidate for the SD. So the SD is 26.71.

(b) Total area of the bars other than 40-70: $0.0013 * (20 - 0) + 0.0108 * (40 - 20) + 0.0077 * (100 - 70) + 0.0015 * (200 - 100) = A$, say.

Height of missing bar: $(1 - A)/(70 - 40)$

Problem 4

Answer:
(a) `ds1a = students.join('SID', sections, 'ID')`

(b) `ds1a.pivot('section', 'year', 'Q4', np.max)`

Problem 5

Answer:

```
def conclusion2(p):
    CUTOFF = .02
    if 0 <= p and p <= CUTOFF:
        return "the data don't support the null"
    elif CUTOFF < p and p <= 1:
        return "the data support the null"
    else:
        return "not a P-value"
```

Problem 6

Answer: $1 - (4/4) * (3/4) * (2/4) * (1/4)$

Problem 7

Answer:

(a) In each cereal box, the chance of each kind of toy is 1/4, regardless of what is in the other boxes.

(b) `0.5*(sum(abs(np.random.multinomial(100, [.25]*4)/100 - 0.25)))`

Problem 8

Answer:

(a) (i) null

(ii) test statistic

(iii) 0.09 or more, where 0.09 comes from a calculation of the TVD, which we used in problem 7b: $0.5 * (|0.27 - 0.25| + |0.2 - 0.25| + |0.21 - 0.25| + |0.32 - 0.25|)$

(b) (i) null

(ii) approximate

(iii) empirical distribution