**Problem 1**

Part 1

calculated mean: 15.04153263 (same as the true mean)

calculated standard deviation: 5.025375677636071, while the true standard deviation is 0.158995850202947. The computed std by numpy is much larger than the true std.

Part 3

Chart, histogram

Description automatically generated

Part 4

The distribution of the means of 1000 samples is close to normal distribution. The mean of the distribution is around the true mean (15.04153263).

Part 5

Chart, histogram

Description automatically generated

Similar to the 25-size one in part 3, the distribution of the means here is also close to normal distribution and its mean is around the true mean. However, in the size-40 one, there are more samples having means close to the true mean (more samples around the center of the graph). Thus, we can get the conclusion that taking larger sized samples will get the sample means closer to the true mean.

Part 6

for the first 25 examples:

Sample mean 14.565156192

Sample standard error: 0.9362242966547015

Confidence interval: (12.632884212773156, 16.497428171226844)

Since the true mean value is 15.04153263, it falls into the 0.95 confidence interval.

**Problem 2**

Part a

The sum of all possible outcomes of these 2 fair dices:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 6 | 7 | 8 | 9 | 10 | 11 | 12 |

Let xi be the outcome of the sum = i, *Ɵ*i be the probability of the sum = i

*Ɵ*2 = 1/36

*Ɵ*3 = 2/36 = 1/18

*Ɵ*4 = 3/36 = 1/12

*Ɵ*5 = 4/36 = 1/9

*Ɵ*6 = 5/36

*Ɵ*7 = 6/36 = 1/6

*Ɵ*8 = 5/36

*Ɵ*9 = 4/36 = 1/9

*Ɵ*10 = 3/36 = 1/12

*Ɵ*11 = 2/26 = 1/18

*Ɵ*12 = 1/36

Part b

The expected value E(x) = 2\* *Ɵ*2 + 3\* *Ɵ*3 +4\* *Ɵ*4 +5\* *Ɵ*5 +6\* *Ɵ*6 +7\* *Ɵ*7 +8\* *Ɵ*8 +9\* *Ɵ*9 +10\* *Ɵ*10 +11\* *Ɵ*11 +12\* *Ɵ*12  = 2\*(1/36) + 3\*(1/18) + 4\*(1/12) + 5\*(1/9) + 6\*(5/36) + 7\*(1/6) + 8\*(5/36) + 9\*(1/9) + 10\*(1/12) + 11\*(1/18) + 12\*(1/36) = 7

Part c

The probability of seeing the outcome of 4 in a single trial is *Ɵ*4 = 1/12

The probability of NOT seeing the outcome of 4 in a single trial is 1-1/12 = 11/12

Thus, the probability of NOT seeing 4 in all 5 trials: (11/12)5 = 161051/248832 ≈ 0.6472

The probability of seeing odd-sum outcomes in a single trial is *Ɵ*3 + *Ɵ*5 + *Ɵ*7 + *Ɵ*9 + *Ɵ*11 = 1/18 + 1/9 + 1/6 + 1/9 + 1/18 = ½

Thus, the probability of seeing odd-sum outcomes in all 5 trials: (1/2)5 = 1/32 = 0.03125

**Problem 3**

(a) Take the integral over its possible values on the interval [a,b]:

So we verified that the uniform distribution is properly normalized.

(b) E(x) =

**Problem 4**

(a) The ML estimate of theta is 0.65.

(b)

Chart, histogram

Description automatically generated

(c) the posterior: Beta(*Ɵ* | 66, 36) =

- calculate the MAP estimate of theta

Take the log of it, we get:

log() + log() + log()

= log() + 65log(Ɵ) + 35log(1-Ɵ)

Take the derivative and set it to 0 to find Ɵ that maximizes the posterior:

0 + = 0 => 65(1-) = 35Ɵ => Ɵ = 0.65

So ƟMAP = 0.65

(or use the equation: ƟMAP =

Here N1 = 65, a = 1, N2 = 35, b = 1,

ƟMAP = = 65/100 = 0.65 )

- calculate the expected value of theta

µ = E[] = =

= = =

Chart, histogram

Description automatically generatedChart, histogram

Description automatically generated

(d)

Chart, histogram

Description automatically generated

the posterior: Beta(*Ɵ* | 69, 37) =

- calculate the MAP estimate of theta

Take the log of it, we get:

log() + log() + log()

= log() + 68log(Ɵ) + 36log(1-Ɵ)

Take the derivative and set it to 0 to find Ɵ that maximizes the posterior:

0 + = 0 => 68(1-) = 36Ɵ => Ɵ = 68/104 = 17/26

So ƟMAP = 17/26

(or use the equation: ƟMAP =

Here N1 = 65, a = 4, N2 = 35, b = 2,

ƟMAP = = 68/104 = 17/26 )

- calculate the expected value of theta

µ = E[] = =

= =

Chart, histogram

Description automatically generated Chart, histogram

Description automatically generated

**Problem 5**

(a)

Chart, histogram

Description automatically generated

(b)

ML estimate of mean is = 3.6377316957999994

ML estimate of unbiased variance is 3.6414284586746737

Chart, histogram

Description automatically generated

**Problem 6**

Part 1

Since the n samples are independent, the likelihood function is

Take the log of it, we get the log likelihood function:

= + +

= n(-) - +

Take its derivative and set it to 0 to find the ML estimate of :

-n – 0 + = 0

=> the ML estimate of

Part 2

(a)

Chart, line chart

Description automatically generated

(b)

Use we derived from part 1, here the ML estimate of 5.24

Chart, line chart

Description automatically generated

(c)

Chart, line chart

Description automatically generated

(d)

Chart, histogram

Description automatically generated

Chart, histogram

Description automatically generated