Check Your Understanding of the Lecture Material

Finger Exercises with Solutions

Instructor: David Dobor

April 15, 2017

While the solutions to these questions are available, you are strongly encouraged to do your best to answer them on your own and to only peek into the solutions if you get stuck or to verify your answers.

Understanding this material is your key to success on your final examination.

Check Your Understanding: 1. CLT

Let X_n be i.i.d. random variables with mean zero and variance σ^2. Let $S_n = X_1 + \cdots + X_n$. Let Φ stand for the standard normal CDF. According to the central limit theorem, and as $n \to \infty$, $\mathbf{P}(S_n \leq 2\sigma\sqrt{n})$ converges to $\Phi(a)$, where:

$a =$

Furthermore,

$\mathbf{P}(S_n \leq 0)$ converges to ...

Check Your Understanding: 2. CLT applicability

Consider the class average in an exam in a few different settings. In all cases, assume that we have a large class consisting of equally well-prepared students. Think about the assumptions behind the central limit theorem, and choose the most appropriate response under the given description of the different settings.

1. Consider the class average in an exam of a fixed difficulty.
   - The CLT can be used to obtain a good normal approximation of the class average
   - The class average is not approximately normal because the student scores are strongly dependent
   - The class average is not approximately normal because the student scores are not identically distributed
2. Consider the class average in an exam that is equally likely to be very easy or very hard.
   - The CLT can be used to obtain a good normal approximation of the class average
   - The class average is not approximately normal because the student scores are strongly dependent
   - The class average is not approximately normal because the student scores are not identically distributed
3. Consider the class average if the class is split into two equal-size sections. One section gets an easy exam and the other section gets a hard exam.
   - The CLT can be used to obtain a good normal approximation of the class average
   - The class average is not approximately normal because the student scores are strongly dependent
   - The class average is not approximately normal because the student scores are not identically distributed
4. Consider the class average if every student is (randomly and independently) given either an easy or a hard exam.
   - The CLT can be used to obtain a good normal approximation of the class average
   - The class average is not approximately normal because the student scores are strongly dependent
   - The class average is not approximately normal because the student scores are not identically distributed

Check Your Understanding: 3. CLT practice

The random variables X_i are i.i.d. with mean 2 and standard deviation equal to 3. Assume that the X_i are nonnegative. Let $S_n = X_1 + \cdots + X_n$. Use the CLT to find good approximations to the following quantities. You may want to refer to the normal table . In parts (a) and (b), give answers with 4 decimal digits.

a) $\mathbf{P}(S_{100} \leq 245) \approx$

b) We let N (a random variable) be the first value of n for which S_n exceeds 119.

   $\mathbf{P}(N > 49) \approx$

c) What is the largest possible value of n for which we have $\mathbf{P}(S_n \leq 128) \approx 0.5$?

   $n =$

Check Your Understanding: 4. CLT for the binomial

Let X be binomial with parameters $n = 49$ and $p = 1/10$.

The meand of X is . . .

The standard deviation of X is . . .

The CLT, together with the 1/2-correction, suggests that

$\mathbf{P}(X = 6) \approx$

You may want to refer to the normal table .

Note: In this case, the CLT may not provide a great approximation. The range of values that X is likely to take is quite narrow, so that its PMF consists of only a few entries of substantial size. But, regardless, we can still calculate what the CLT suggests.

Table entry for z is the area under the standard normal curve to the left of z.

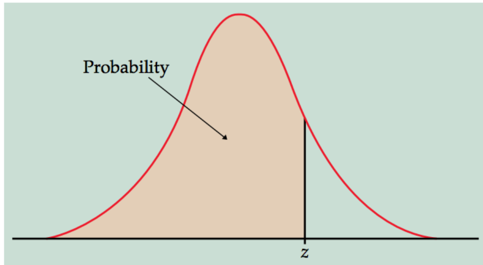

TABLE A Standard normal probabilities (*continued*)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Answers

QUESTION 1

We have

$$\lim_{n\to\infty} \mathbf{P}(S_n \le 2\sigma\sqrt{n}) = \lim_{n\to\infty} \mathbf{P}\left(\frac{S_n - 0}{\sigma\sqrt{n}} \le 2\right) = \Phi(2).$$

Similarly,

$$\lim_{n\to\infty} \mathbf{P}(S_n \le 0) = \lim_{n\to\infty} \mathbf{P}\left(\frac{S_n - 0}{\sigma\sqrt{n}} \le 0\right) = \Phi(0) = \frac{1}{2}.$$

QUESTION 2

1. Since students are equally well-prepared and the difficulty level is fixed, the only randomness in a student's score comes from luck or accidental mistakes of that student. It is then plausible to assume that each student's score will be an independent random variable drawn from the same distribution, and the CLT applies.
2. Here, the score of each student depends strongly on the difficulty level of the exam, which is random but common for all students. This creates a strong dependence between the student scores, and the CLT does not apply.
3. This is more subtle. The scores of the different students are not identically distributed. However, let Y_i be the score of the ith student from the first section and let Z_i be the score of the ith student in the second section. The class average is the average of the random variables $(Y_i + Z_i)/2$. Under our assumptions, these latter random variables can be modeled as i.i.d., and the CLT applies.
4. Unlike part (2), here the student scores are i.i.d., and the CLT applies.

QUESTION 3

We will use Z_n to refer to the standardized random variable

$$\frac{S_n - 2n}{3\sqrt{n}}.$$

a) We have

$$\mathbf{P}(S_{100} \le 245) = \mathbf{P}\left(\frac{S_{100} - 2\cdot 100}{3\cdot\sqrt{100}} \le \frac{245 - 2\cdot 100}{3\cdot\sqrt{100}}\right) = \mathbf{P}(Z_n \le 1.5) \approx 0.9332.$$

b) The event $N > 49$ is the same as the event $S_{49} \le 119$. Its probability is

$$\mathbf{P}(S_{49} \le 119) = \mathbf{P}\left(\frac{S_{49} - 2\cdot 49}{3\cdot\sqrt{49}} \le \frac{119 - 2\cdot 49}{3\cdot\sqrt{49}}\right) = \mathbf{P}(Z_n \le 1) \approx 0.8413.$$

c) We want n such that

$$0.5 \approx \mathbf{P}(S_n \leq 128) = \mathbf{P}\left(\frac{S_n - 2n}{3\sqrt{n}} \leq \frac{128 - 2n}{3\sqrt{n}}\right) = \Phi\left(\frac{128 - 2n}{3\sqrt{n}}\right).$$

But since $0.5 = \Phi(0)$, we must have $(128 - 2n)/(3\sqrt{n}) = 0$, so that $n = 128/2 = 64$.

A faster way to see the answer is to note that since the normal is symmetric around its mean, the relation $\mathbf{P}(S_n \leq 128) \approx 0.50$ tells us that 128 should be equal to the mean, $2n$, of S_n.

QUESTION 4

We have $\mathbf{E}[X] = np = 4.9$ and

$$\text{var}(X) = np(1-p) = 49 \cdot \frac{1}{10} \cdot \frac{9}{10} = \frac{49 \cdot 9}{10^2},$$

so that the standard deviation of X is $21/10 = 2.1$.

The standardized version of X is $(X - 4.9)/2.1$. Thus,

$$\begin{aligned}\mathbf{P}(X = 6) = \mathbf{P}(5.5 < X < 6.5) &= \mathbf{P}\left(\frac{5.5 - 4.9}{2.1} \leq \frac{X - 4.9}{2.1} \leq \frac{6.5 - 4.9}{2.1}\right) \\ &\approx \Phi(0.76) - \Phi(0.29) \approx 0.7764 - 0.6141 = 0.1623.\end{aligned}$$

For comparison, the answer calculated by using the binomial PMF directly is

$$\mathbf{P}(X = 6) = \binom{49}{6}(0.1)^6(0.9)^{49-6} \approx 0.1507.$$