Homework 1. Solutions.

1. Problem 13. Each step is a random variable

\[ X_i = \begin{cases} 
50 & \text{with probability } \frac{1}{2}, \\
-50 & \text{with probability } \frac{1}{2}.
\end{cases} \]

After 60 minutes the drunkard takes 60 steps \( S_{60} = X_1 + \ldots + X_{60} \). Note

\[ ES_{60} = 60\mu = 60(50\cdot \frac{1}{2} - 50\cdot \frac{1}{2}) = 0, \quad \text{Var} S_{60} = ES_{60}^2 = 60(50^2\cdot \frac{1}{2} + (-50)^2\cdot \frac{1}{2}) = 150000. \]

By the CLT the approximate probability distribution of \( S_{60} \) is normal with mean 0 and variance 150000. The drunkard is most likely to be at 0, where he has started.

2. Problem 15. Each bet is a random variable \( X_i = \begin{cases} 
5 & \text{with probability } \frac{1}{2}, \\
-5 & \text{with probability } \frac{1}{2}.
\end{cases} \)

After 50 games you will have \( S_{50} = X_1 + \ldots + X_{50} \) dollars. Note

\[ ES_{50} = 50\mu = 50(5\cdot \frac{1}{2} - 5\cdot \frac{1}{2}) = 0, \quad \text{Var} S_{50} = ES_{50}^2 = 50(5^2\cdot \frac{1}{2} + (-5)^2\cdot \frac{1}{2}) = 1250. \]

By the CLT

\[ P(S_{50} < -75) = P \left( \frac{S_{50} - 0}{\sqrt{1250}} < \frac{-75 - 0}{\sqrt{1250}} \right) \approx P(Z < -2.12) = 1 - P(Z < 2.12) = 1 - 0.9830 = 0.017. \]

3. Problem 16. Note

\[ EX_1 = \int_0^1 x^2 \, dx = \int_0^1 2x^2 \, dx = \frac{2x^3}{3} \bigg|_0^1 = \frac{2}{3}. \]

\[ E(X_1^2) = \int_0^1 x^4 \, dx = \int_0^1 2x^3 \, dx = \frac{2x^4}{4} \bigg|_0^1 = \frac{1}{2}. \]

\[ Var(X_1) = E(X_1^2) - (EX_1)^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}. \]

\[ P(S \leq 10) = P \left( \frac{S - 20\cdot \frac{2}{3}}{\sqrt{20\cdot \frac{1}{18}}} \leq \frac{10 - 20\cdot \frac{2}{3}}{\sqrt{20\cdot \frac{1}{18}}} \right) \approx P(Z \leq -3.16) = 1 - P(Z \leq 3.16) = 0.0008. \]

4. Problem 18. We can use the CLT with \( \mu = 15 \) and \( \sigma = 10 \):

\[ P(S_{100} > 1700) = P \left( \frac{S_{100} - 1500}{\sqrt{10000}} > \frac{1700 - 1500}{\sqrt{10000}} \right) \approx P(Z > 2) = 1 - P(Z \leq 2) = 1 - 0.9772 = 0.0228. \]

5. Problem 26. A successful shot can be modelled as a Bernoulli random variable with \( p = 0.3 \). Then \( S_{25} \) is a binomial random variable with parameters 25 and 0.3. It has the mean \( np = 25\cdot 0.3 = 7.5 \) and the variance \( np(1 - p) = 5.25 \). For each number \( a \) we can use the CLT to approximate

\[ P(S \leq a) = P \left( \frac{S - 7.5}{\sqrt{5.25}} \leq \frac{a - 7.5}{\sqrt{5.25}} \right) \approx P \left( Z \leq \frac{a + 0.5 - 7.5}{2.29} \right), \]

where 0.5 is a correction coefficient. Then
<table>
<thead>
<tr>
<th>a</th>
<th>exact $P(S \leq a)$</th>
<th>approximate $P(S \leq a)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.1935</td>
<td>$P(Z \leq -0.87) = 0.1922$</td>
</tr>
<tr>
<td>7</td>
<td>0.5118</td>
<td>$P(Z \leq 0) = 0.5$</td>
</tr>
<tr>
<td>9</td>
<td>0.8106</td>
<td>$P(Z \leq 0.87) = 0.8078$</td>
</tr>
<tr>
<td>11</td>
<td>0.9558</td>
<td>$P(Z \leq 1.75) = 0.9599$</td>
</tr>
</tbody>
</table>

6. Problem 19. We can use for example Matlab and generate $U_1, ..., U_n$ from uniform distribution on $[0, 1]$.

(a) $\frac{1}{n} \sum_{i=1}^{n} \cos(2\pi U_i)$ is 0.0597 for $n = 100$ and -0.0245 for $n = 1000$. The exact answer is 0.

(b) $\frac{1}{n} \sum_{i=1}^{n} \cos(2\pi U_i)$ is 0.1668 for $n = 100$ and 0.2293 for $n = 1000$. The exact answer is very difficult to obtain analytically. But for $n = 10^8$ the MC approximation is 0.2443 ± 0.0001.

7. This is a normal approximation to binomial distribution for $n = 80, p = 0.1$.

(a) One can use

$$P(S > 2) = P((S - np)/\sqrt{np(1-p)} > (2 - 8)/\sqrt{7.2}) \approx P(Z > -2.24);$$

also one can use $P(S > 2.5) \approx P(Z > -2.05)$.

(b) Use

$$P(S = 2) = P((1.5-8)/\sqrt{7.2} \leq (S - np)/\sqrt{np(1-p)} \leq (2.5-8)/\sqrt{7.2}) \approx P(-2.42 < Z < -2.05).$$

8. (a) The width of 2 means $|\bar{X} - \mu| \leq 1$. So $P(|\bar{X} - \mu| \leq 1) = 0.95$, then $P(|\bar{X} - \mu|/(\sigma/\sqrt{n}) \leq 1/(\sigma/\sqrt{n})) = 0.95$ which is approximately $P(|Z| \leq \sqrt{n}/\sigma) \approx 0.95$. Then $\sqrt{n}/10 = 1.96$ and $n = 385$ after rounding up.

(b) If the random procedure is repeated many times, we can guarantee that 95% of the different confidence intervals would cover the true mean $\mu_X$. 

6.