1. A biased coin, \( P(H) = 0.3 \), is tossed 100 times. Let \( X \) be the number of heads that result.
   (a) Set up a sum for the exact probability that \( X \leq 20 \), using the binomial distribution, but do not attempt to add up the sum.
   (b) Use a normal approximation to compute the probability that \( X \leq 20 \).

2. A sequence of random points \( X_1, X_2, \ldots, X_{500} \) are independently chosen from interval \([0, 2]\), where each \( X_i \) has density (pdf) given by
   \[
   f(x) = \frac{x}{2}.
   \]
   (a) Set \( \mu = E(X_1) \). What is \( \mu \)?
   (b) Set \( \sigma = \sigma(X_1) \). What is \( \sigma \)?
   (c) Let \( \bar{x} \) be the average of the values \( X_1, X_2, \ldots, X_{500} \). What is the probability that \( |\bar{x} - \mu| < 0.05 \).
   \[\text{Hint: By the central limit theorem, you can use a normal approximation.}\]

3. A biased coin, \( P(H) = p \), is tossed 40 times, resulting in 12 heads and 28 tails.
   (a) Estimate \( p \), and find an 80% confidence interval for this estimate.
   (b) What is the margin of error in part (a)?
   (c) Find a 99% confidence interval for your estimate for \( p \) in part (a).
   (d) Suppose you can toss the coin as many times as you want in order to estimate \( p \). What is the minimum number of times one should toss the coin in order to estimate \( p \) accurately to two decimal places with 95% confidence?

4. An article claims that 63% of all scientific studies use statistics incorrectly. The article gives this percentage along with a margin of error of 2%. Assuming that the assertion was made with 95% confidence, how many scientific studies were actually checked for correctness by the authors of this article?

\[\text{Solutions on the next page} \rightarrow\]
Solutions:

1. (a) \( X \) has a binomial distribution with \( p = 0.3 \) and \( n = 100 \), so that \( P(0 \leq X \leq 20) = \sum_{k=0}^{20} \binom{100}{k}(0.3)^k(0.7)^{100-k} \).

(b) We have \( np = 30 \) and \( np(1-p) = 21 \). A normal approximation (with continuity correction) yields

\[
P(X \leq 20.5) = P \left( \frac{X - 30}{\sqrt{21}} \leq \frac{20.5 - 30}{\sqrt{21}} \right) \approx P \left( Z \leq \frac{-9.5}{\sqrt{21}} \right) = F_Z(-0.2073) = 0.0192 \approx 1.92%\]

Remark: In case you were wondering, the (correct) probability from the sum in part (a) is 0.01646 or about 1.65%.

2. (a) \( \mu = E(X_1) = \int_0^x \frac{x}{2} \, dx = \frac{x^2}{4} \)

(b) Since \( E(X_i^2) = \int_0^x x^2 \, dx = 2 \), we have \( \sigma = \frac{x^2}{2} \)

(c) Note that \( n = 500 \). We then have

\[
P(|\bar{x} - \mu| < 0.05) = P \left( \frac{|\bar{x} - \mu|}{\sigma / \sqrt{n}} < 0.05 \right) \approx P \left( Z < \frac{0.05}{\sqrt{\frac{2}{500}}} \right) = F_Z(2.37) = 0.9911 \approx 99.1\%
\]

3. (a) Estimate \( p \) by \( \bar{x} = \frac{12}{40} = 0.3 \).

At 80% confidence (\( C = 0.8 \)), we choose \( z^* \) so that \( 1 - F_Z(z^*) = \frac{1-C}{2} = 0.1 \) or \( F_Z(z^*) = 0.9 \). This occurs when \( z^* = 1.28 \). We then have

\[
0.8 \approx P \left( |p - \bar{x}| < \frac{z^* \sqrt{\bar{x}(1-\bar{x})}}{\sqrt{n}} \right) = P \left( |p - 0.3| < \frac{1.28 \sqrt{(0.3)(0.7)}}{\sqrt{400}} \right) = P \left( |p - 0.3| < 0.093 \right).
\]

So we are 80% confident that \( 0.207 \leq p \leq 0.393 \)

(b) From part (a) we have a margin of error \( m \) is \( m = \frac{z^* \sqrt{\bar{x}(1-\bar{x})}}{\sqrt{n}} = \frac{1.28 \sqrt{0.37}}{\sqrt{400}} \approx 0.093 \).

However, in experimental design the value of \( \bar{x} \), and therefore of \( \sqrt{\bar{x}(1-\bar{x})} \) is not known in advance, so the more conservative formula

\[
m = \frac{z^*}{2 \sqrt{n}} = \frac{1.28}{2 \sqrt{400}} = 0.101
\]

may be used.

(c) For 99% confidence, we must proceed as in part (a), using \( z^* \) so that \( 1 - F_Z(z^*) = \frac{1-0.99}{2} = 0.005 \) or \( F_Z(z^*) = 0.995 \). This occurs when \( z^* = 2.58 \). We then have

\[
0.99 \approx P \left( |p - 0.3| < \frac{2.58 \sqrt{(0.3)(0.7)}}{\sqrt{40}} \right) = P \left( |p - 0.3| < 0.187 \right).
\]

So we are 99% confident that \( 0.113 \leq p \leq 0.487 \)

(d) For a 95% confidence estimate we use \( z^* = 1.96 \). To achieve 0.01 accuracy, we need

\[
m = \frac{z^*}{2 \sqrt{n}} = \frac{1.96}{2 \sqrt{n}} = 0.01,
\]

so that \( n = 9604 \) is the minimum number of coin tosses required.

4. If the precise margin of error (at 95% confidence) is 2%, where \( \bar{x} = 0.63 \)

\[
0.02 = \frac{1.96 \sqrt{0.63 \cdot 0.37}}{\sqrt{n}},
\]

so that \( n = 2239 \).

However, the experimenters did not know the value \( \bar{x} = 0.63 \) at the start of their experiment. If they designed the experiment to achieve \( m = 0.02 \) at 95% confidence, they would choose \( n \) via the more conservative formula

\[
0.02 = \frac{1.96}{2 \sqrt{n}} \text{ or, equivalently, } n = \left( \frac{1.96}{2 \cdot 0.02} \right)^2,
\]

so that \( n = 2401 \).