과제#2

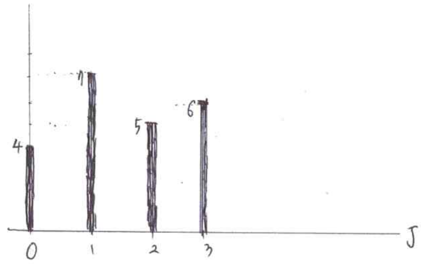
20213064_김종민

Problem 3.2 Repeat Problem 3.1 for the discrete random variable J that takes on the values

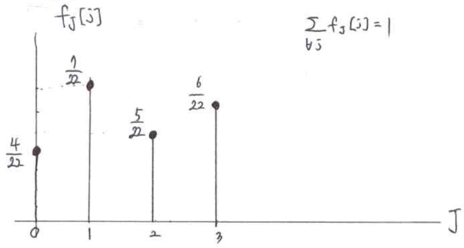
[1, 3, 0, 2, 1, 2, 0, 3, 1, 1, 3, 1, 0, 1, 2, 3, 0, 2, 1, 3, 3, 2]

↳ 0이 4개, 1이 7개, 2가 5개, 3이 6개 총 22개 숫자.

(a)



(b)



Problem 4.11

(a) Make a plot of the probability that Roberto does not ask Claudia for a date given that he has not asked her by the end of l^{th} day $1 \le l \le 5$.

↳ 우선 사건을 정의합니다.

B_l: Roberto has not asked Claudia by the end of the l^{th} day. $1 \le l \le 5$.

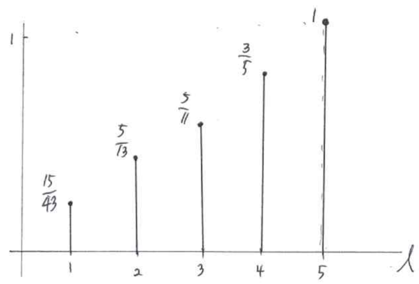
$Pr[B_l] = \frac{1}{3} + \sum_{k=l+1}^{5} \frac{2}{3} \cdot \frac{k}{15} = 1 - \sum_{k=1}^{l} \frac{2}{3} \cdot \frac{k}{15}$

$Pr[D^c | B_l] = \frac{Pr[D^c \cdot B_l]}{Pr[B_l]}$

$Pr[D] = \frac{2}{3}$ 이므로 $Pr[D^c] = \frac{1}{3}$

$Pr[B_1] = \frac{43}{45}$, $Pr[B_2] = \frac{39}{45}$, $Pr[B_3] = \frac{33}{45}$

$Pr[B_4] = \frac{25}{45}$, $Pr[B_5] = \frac{15}{45}$



(b) Sketch the conditional PMF for K given that Roberto asks Claudia out but has not done so by the end of the second day. Given this situation, what is the probability that Roberto asks her out

(i) on the third day $(k=3)$?

(ii) on the fifth day $(k=5)$?

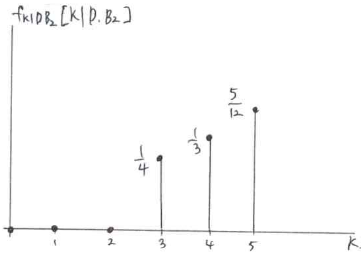
↳ Roberto가 2번째 날까지는 데이트 신청을 하지 않고 3, 4, 5번째 날에 데이트 신청을 할 확률은 $\frac{6}{45}+\frac{8}{45}+\frac{10}{45} = \frac{24}{45}$ 이다.

3번째 날에 데이트 신청할 확률은 $(6/45)/(24/45) = \frac{1}{4}$

4번째날에 데이트 신청할 확률은 $(8/45)/(24/45) = \frac{1}{3}$

5번째날에 데이트 신청할 확률은 $(10/45)/(24/45) = \frac{5}{12}$

이를 바탕으로 PMF는 그리면 오른쪽과 같다.



(a)에서 2번째 날까지 Roberto가 Claudia에게 묻지 않을 확률 $(Pr[B_2])$가 $\frac{39}{45}$ 라고 했습니다.

↳ (i) $Pr[K=3, D \mid B_2] = (6/45)/(39/45) = \underline{\frac{2}{13}}$

↳ (ii) $Pr[K=5, D \mid B_2] = (10/45)/(39/45) = \underline{\frac{10}{39}}$

(c) By the middle of the third day of the week Roberto has not asked Claudia for a date; but Rolf decides to ask her. Will she accept Rolf or not?

↳ Rolf와 같이 있을 때 Claudia의 emotional state (α)는 4이므로

Rolf의 데이트를 받아줄 기댓값은 $E(\alpha) = 4$ 이다.

한편, Rolf의 데이트를 거절하는 기댓값은 Roberto가 데이트 신청을 할지 안 할지로 결정된다.

Roberto와 데이트 할 때의 α는 10, 토요일에 데이트 하지 못할 때의 α는 -5이다.

앞서 $Pr[D^c \mid B_2] = \frac{5}{13}$, $Pr[D \mid B_2] = \frac{8}{13}$ 임을 확인했다.

Roberto에게 No라고 할 기댓값은 $Pr[D \mid B_2] \times 10 + Pr[D \mid B_2] \times (-5)$

$$= \frac{8}{13} \times 10 + \frac{5}{13} \times (-5) = \frac{55}{13} = 4.2307....$$

$$\therefore E(\alpha) = 4.2307$$

기댓값을 비교했을때 Claudia는 Rolf를 거절할 것이다.

(d) Rolf has not studied statistics (not even probability), and thinks that his chance for a date with Claudia will be better if he asks her earlier in the week. Is he right or wrong?

↳ (c)에서 구한 것과 같이 기댓값을 구하여 비교한다.

$l=1$일 때 $E(l)$는 $Pr[D|B_1] \times 10 + Pr[D^c|B_1] \times (-5)$와 같다.

$$Pr[D|B_1] \times 10 + Pr[D^c|B_1] \times (-5) = \frac{28}{43} \times 10 + \frac{15}{43} \times (-5) = \frac{205}{43} = 4.767...$$

$E(l) = 4.767$ 이다.

(c)에서 구한 기댓값 보다 큼을 확인할 수 있고, 이는 거절당할 기댓값이 커짐을 의미한다.

∴ Rolf의 생각은 틀렸다.

(e) What is the optimal strategy for Rolf (i.e., when should he ask Claudia) in order to maximize the probability that Claudia will accept if he asks her for a Saturday night date?

↳ $l=1, 2, 3, 4, 5$의 기댓값을 전부 비교해봅니다.

$l=1$일 때, $Pr[D|B_1] \times 10 + Pr[D^c|B_1] \times (-5) = \frac{28}{43} \times 10 + \frac{15}{43} \times (-5) = \frac{205}{43} = 4.767$

$l=2$일 때, $Pr[D|B_2] \times 10 + Pr[D^c|B_2] \times (-5) = \frac{8}{13} \times 10 + \frac{5}{13} \times (-5) = \frac{55}{13} = 4.231$

$l=3$일 때, $Pr[D|B_3] \times 10 + Pr[D^c|B_3] \times (-5) = \frac{6}{11} \times 10 + \frac{5}{11} \times (-5) = \frac{35}{11} \doteq 3.182$

$l=4$일 때, $Pr[D|B_4] \times 10 + Pr[D^c|B_4] \times (-5) = \frac{2}{5} \times 10 + \frac{3}{5} \times (-5) = 1$

$l=5$일 때, $Pr[D|B_5] \times 10 + Pr[D^c|B_5] \times (-5) = 0 \times 10 + 1 \times (-5) = -5$

위의 결과에서 확인할 수 있듯이 3번째 날까지 Roberto가 데이트 신청을 하지 않았을 때 처음으로 $E(l)$가 4보다 작아진다.

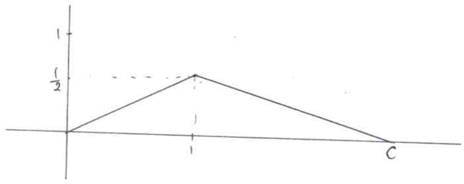
∴ 3번째 날 이후인 4번째나 5번째 날에 Rolf는 데이트 신청을 해야 한다.

$\boxed{\text{Problem 4.12}}$ The probability density function of a random variable X is given by

$$f_X(x) = \begin{cases} \frac{1}{2}x & (0 \le x \le 1) \\ \frac{1}{6}(4-x) & (1 < x < c) \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the constant C

↳ PDF는 x축과 함수로 둘러싸인 넓이가 1이다.



$\frac{1}{2}x$ $(0 \le x \le 1)$의 넓이가 $\frac{1}{4}$이므로 $\frac{1}{6}(4-x)$ $(1<x<c)$의 넓이가 $\frac{3}{4}$이어야 한다.

$$(c-1)\cdot\left(\frac{1}{2} + \frac{1}{6}(4-c)\right) \times \frac{1}{2} = \frac{3}{4} \rightsquigarrow c^2 - 8c + 16 = 0 \rightsquigarrow c = 4$$

$\therefore$ constant $C = 4$

(b) calculate the expectation : $E[3+2X]$.

↳ PDF의 기댓값은 오른쪽과 같이 정의됩니다. $E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$.

$$\int_{-\infty}^{\infty} x \cdot f(x)\,dx = \int_0^4 x f(x)\,dx = \int_0^1 x \cdot \frac{1}{2}x\,dx + \int_1^4 x \cdot \frac{1}{6}(4-x)\,dx = \frac{1}{2}\int_0^1 x^2\,dx + \frac{1}{6}\int_1^4 (4x - x^2)\,dx$$

$$= \frac{1}{2}\left[\frac{1}{3}x^3\right]_0^1 + \frac{1}{6}\left[2x^2 - \frac{x^3}{3}\right]_1^4 = \frac{1}{6} + \frac{3}{2} = \frac{5}{3}$$

$E(X) = \frac{5}{3}$이고 기댓값의 성질에 의해 $E(3+2X) = \frac{19}{3}$ 라는 값을 구할 수 있다.

$\therefore E[3+2X] = \frac{19}{3}$

Problem 6.1 The magnitude of the voltage V across a component in an electronic circuit has a mean value of 0.45 volts. Given only this information, find a bound on the probability that $V \geq 1.35$.

↳ 기댓값만 주어져있기 때문에 마르코프 부등식을 활용해야 한다.

$$E(X) = \int_0^\infty x f(x)\,dx = \int_0^a x f(x)\,dx + \int_a^\infty x f(x)\,dx \geq \int_a^\infty x f(x)\,dx \geq a \int_a^\infty f(x)\,dx = a\,P_r[x \geq a]$$

$$\therefore P_r[x \geq a] \leq \frac{E(X)}{a}$$

위의 식을 이용해서 $V \geq 1.35$ 이상이기 때문에 $a = 1.35$ 평균이 0.45이므로

$$P_r[V \geq 1.35] \leq \frac{0.45}{1.35} = \frac{1}{3}$$

$\therefore$ 확률의 상한은 $\frac{1}{3}$이다.

problem 4.43 The probability density function of a Laplace random variable is given by
$$f_X(x) = \frac{\alpha}{2} e^{-\alpha|x-\mu|}, \quad -\infty < x < \infty$$

(a) Find the mean m_X, second moment, and variance σ_X^2.

↳ 구하고자 하는 것들은 적률 생성 함수를 구하여 구할 수 있다.

$$M_X(s) = E[e^{sX}] = \int_{-\infty}^{\infty} f_X(x) e^{sX}\,dX$$

$$M_X(s) = \int_{-\infty}^{\infty} \frac{\alpha}{2} e^{-\alpha|x-\mu|} \cdot e^{sX}\,dX = \int_{-\infty}^{\mu} \frac{\alpha}{2} e^{\alpha(x-\mu)} \cdot e^{sX}\,dX + \int_{\mu}^{\infty} \frac{\alpha}{2} e^{-\alpha(x-\mu)} \cdot e^{sX}\,dX$$

위의 식을 간단히 하면 $M_X(s) = \dfrac{\alpha^2 \cdot e^{\mu s}}{\alpha^2 - s^2}$ 가 된다.

1차 적률이 곧 mean m_X이므로 $M_X(s)' = \alpha^2 \mu e^{\mu s}(\alpha^2 - s^2)^{-1} + \alpha^2 e^{\mu s}(2s)(\alpha^2 - s^2)^{-2}$

$$M_X(0)' = \mu \qquad \therefore \text{mean } m_X = \mu.$$

2차 적률은 $M_X(s)''$ 이다. $M_X(s)'$를 한번더 미분하면

$$M_X(s)'' = \alpha^2 \mu \cdot \mu \cdot e^{\mu s}(\alpha^2 - s^2)^{-1} + \alpha^2 \mu e^{\mu s}(2s) \cdot (\alpha^2 - s^2)^{-2} + \alpha^2 \mu e^{\mu s}(2s)(\alpha^2 - s^2)^{-2} + \alpha^2 e^{\mu s} \cdot 2(\alpha^2 - s^2)^{-2} + \alpha^2 e^{\mu s}(2s)(4s)(\alpha^2 - s^2)^{-3}$$

$$M_X(0)'' = \mu^2 + 0 + 0 + \frac{2}{\alpha^2} + 0 \qquad \therefore \text{second moment} = \mu^2 + \frac{2}{\alpha^2}$$

마지막으로 분산은 $E(X^2) - (E(X))^2 = V(X)$ 이다. 위에서 mean m_X $(E(X))$ 와 second moment $(E(X^2))$를 구했으니 대입하면 $V(X) = \mu^2 + \frac{2}{\alpha^2} - (\mu)^2 = \frac{2}{\alpha^2}$

$$\therefore \sigma_X^2 = \frac{2}{\alpha^2}$$

(b) Determine the following probabilities for $\alpha = 4$ and $\mu = 1$.

(i) $Pr[\,|X - m_x| > 2\sigma_x\,]$

↳ $f_x(x) = \frac{4}{2} e^{-4|x-1|} = 2e^{-4|x-1|}$

$E(X) = m_x = \mu = 1$, $E(X^2) = \mu^2 + \frac{2}{\alpha^2} = \frac{9}{8}$, $V(X) = \sigma^2(X) = \frac{1}{8}$, $\sigma(X) = \frac{\sqrt{2}}{4}$

$Pr[\,|X - m_x| > 2\sigma_x] = Pr[X > m_x + 2\sigma_x] + Pr[X < m_x - 2\sigma_x] = Pr[X > 1 + \frac{\sqrt{2}}{2}] + Pr[X < 1 - \frac{\sqrt{2}}{2}]$

$$\int_{1+\frac{\sqrt{2}}{2}}^{\infty} 2e^{-4(x-1)}\,dx + \int_{-\infty}^{1-\frac{\sqrt{2}}{2}} 2e^{4(x-1)}\,dx = 2\left[\frac{e^{-4(x-1)}}{-4}\right]_{1+\frac{\sqrt{2}}{2}}^{\infty} + 2\left[\frac{e^{4(x-1)}}{4}\right]_{-\infty}^{1-\frac{\sqrt{2}}{2}}$$

$= e^{-2\sqrt{2}}$

$\therefore Pr[\,|X - m_x| > 2\sigma_x] = 0.05911$

(ii) $Pr[\,|X - m_x| < \sigma_x]$

↳ $Pr[\,|X - m_x| < \sigma_x] = Pr[-\sigma_x + m_x < X < \sigma_x + m_x] = Pr[-\frac{\sqrt{2}}{4} + 1 < X < \frac{\sqrt{2}}{4} + 1]$.

$$\int_{1-\frac{\sqrt{2}}{4}}^{1} 2e^{4(x-1)}\,dx + \int_{1}^{1+\frac{\sqrt{2}}{4}} 2e^{-4(x-1)}\,dx = 2\left[\frac{e^{4(x-1)}}{4}\right]_{1-\frac{\sqrt{2}}{4}}^{1} + 2\left[\frac{e^{-4(x-1)}}{-4}\right]_{1}^{1+\frac{\sqrt{2}}{4}}$$

$= 1 - e^{-\sqrt{2}}$

$\therefore Pr[\,|X - m_x| < \sigma_x] = 0.7569$

Problem 4.49 The moment generating function of a continuous random variable X is given to be $M_x(s) = \frac{3}{1-2s}$.

Find the second moment.

↳ n차 적률 계산 (n^{th} moment) → $E(X^n) = \left.\frac{d^n M_x(s)}{ds^n}\right|_{s=0}$.

즉 $M_x(s) = \frac{3}{1-2s}$ 을 두번 미분하고 $s=0$을 대입하면 second moment($E(X^2)$)을 구할 수 있다.

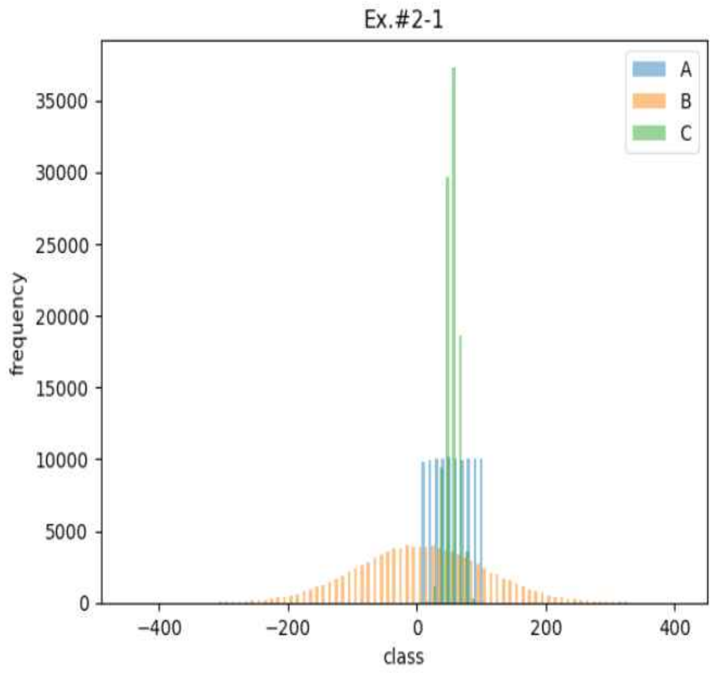
$M_x(s)' = 6(1-2s)^{-2}$

$M_x(s)'' = 24(1-2s)^{-3}$

$M_x(0)'' = 24$

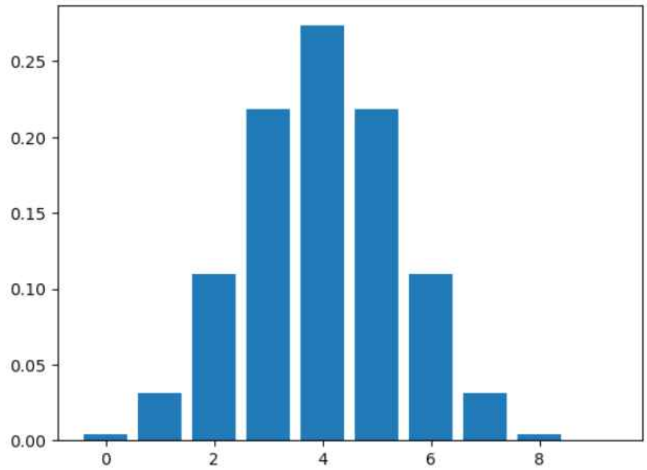
$\therefore$ second moment $E(X^2) = 24$

In [17]:
```python
# 실습 3. 히스토그램 생성(NumPy)
import numpy as np
import matplotlib.pyplot as plt
n_sample = 100000
A = np.random.rand(n_sample)
B = np.random.randn(n_sample)
C = 0.5 + np.random.randn(n_sample) * 0.1
sampleA, sampleB, sampleC = A*100, B*100, C*100
classA = np.arange(min(sampleA), max(sampleA) + 1, 10)
classB = np.arange(min(sampleB), max(sampleB) + 1, 10)
classC = np.arange(min(sampleC), max(sampleC) + 1, 10)

counts1, bins1 = np.histogram(sampleA, classA)
counts2, bins2 = np.histogram(sampleB, classB)
counts3, bins3 = np.histogram(sampleC, classC)

plt.bar(bins1[1:], counts1, alpha = 0.5, width = 5, label = 'A')
plt.bar(bins2[1:], counts2, alpha = 0.5, width = 5, label = 'B')
plt.bar(bins3[1:], counts3, alpha = 0.5, width = 5, label = 'C')

plt.legend()
plt.title("Ex.#2-1")
plt.xlabel("class")
plt.ylabel("frequency")
plt.show()

print("A (mean) : ",np.mean(A))
print("A (std) :", np.std(A))
print("B (mean) : ",np.mean(B))
print("B (std) :", np.std(B))
print("C (mean) : ",np.mean(C))
print("C (std) :", np.std(C))
```

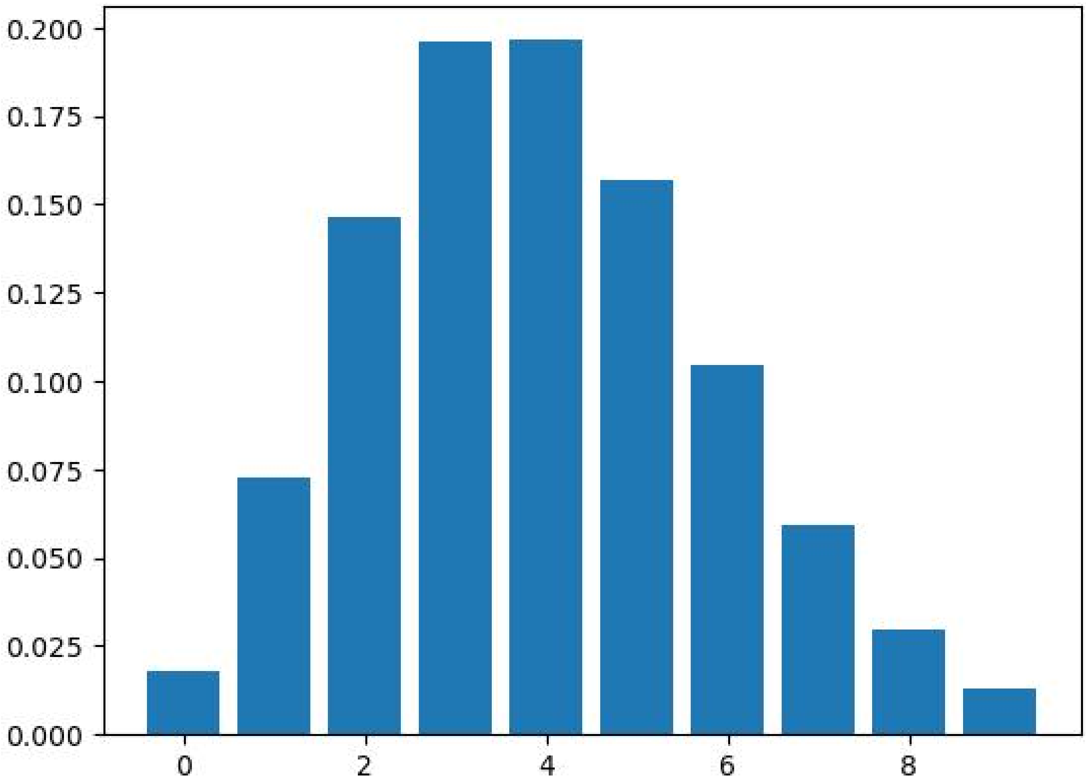



```
A (mean) :  0.5006570026302329
A (std) : 0.2881654420736816
B (mean) :  0.0035647828930354813
B (std) : 1.0030496629149745
C (mean) :  0.4998845647132035
C (std) : 0.0995742560111345
```

In [19]:
```
# 실습 4. Poisson 확률분포
import scipy.stats as st
p = 1/2
x = np.arange(0, 10)
y = st.binom.pmf(x, 8, p)
plt.bar(x, y)
plt.show()
```

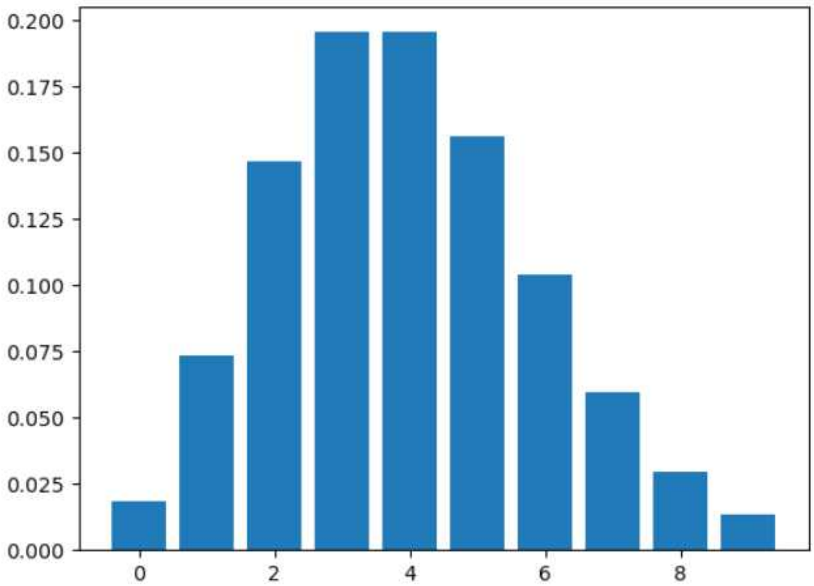


In [21]:
```
p = 1/100
x = np.arange(0,10)
y = st.binom.pmf(x, 400, p)
plt.bar(x, y)
plt.show()
```



In [23]:

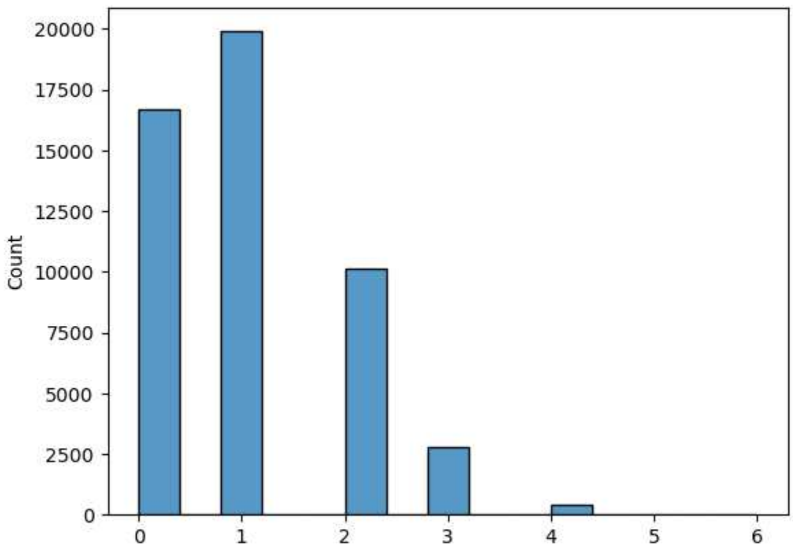
```
pn = 4
x = np.arange(0, 10)
y = st.poisson.pmf(x, pn)
plt.bar(x, y)
plt.show()
```



In [104]:

```
# 실습 5. Binomial 분포에 대한 중심극한정리
# N = 1, n = 6, P = 1/6
import seaborn as sns
N =  1
p = 1/ 6
x_sums = np.zeros(50000)
for k in range(50000):
    x = st.binom.rvs(6, p, size = N)
    x_sums[k] = np.sum(x)
x_sum_mean = np.mean(x_sums)
x_sum_std = np.std(x_sums)
print(x_sum_mean, x_sum_std ** 2)
sns.histplot(x_sums, bins = 15, kde = False)
plt.show()
```

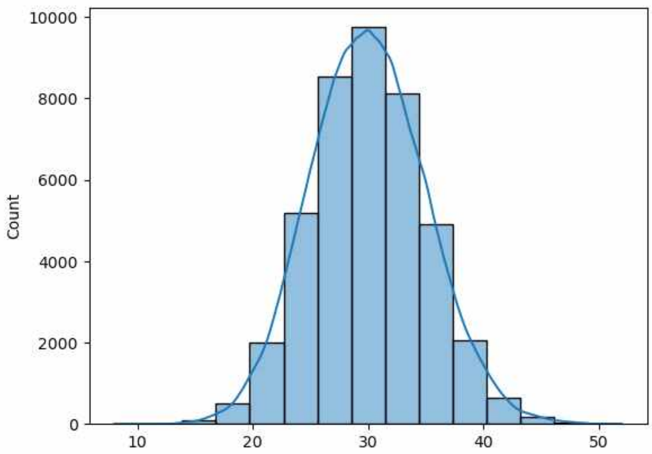
```
1.00914 0.8480164603999998
```



In [102]:

```python
# 실습 5. Binomial 분포에 대한 중심극한정리
# N = 30, n = 6, p = 1/6
import seaborn as sns
N = 30
p = 1/ 6
x_sums = np.zeros(42000)
for k in range(42000):
        x = st.binom.rvs(6, p, size = N)
        x_sums[k] = np.sum(x)
x_sum_mean = np.mean(x_sums)
x_sum_std = np.std(x_sums)
print(x_sum_mean, x_sum_std ** 2)
sns.histplot(x_sums, bins = 15, kde = True)
plt.show()
```

30.032761904761905 25.03864094331066



In [101]:

```python
# 실습 5. Binomial 분포에 대한 중심극한정리
# N = 6000, n = 6, p = 1/6
import seaborn as sns
N = 6000
p = 1/ 6
x_sums = np.zeros(42000)
for k in range(42000):
        x = st.binom.rvs(6, p, size = N)
        x_sums[k] = np.sum(x)
x_sum_mean = np.mean(x_sums)
x_sum_std = np.std(x_sums)
print(x_sum_mean, x_sum_std ** 2)
sns.histplot(x_sums, bins = 15, kde = True)
plt.show()
```

5999.628214285714 5033.661037287415