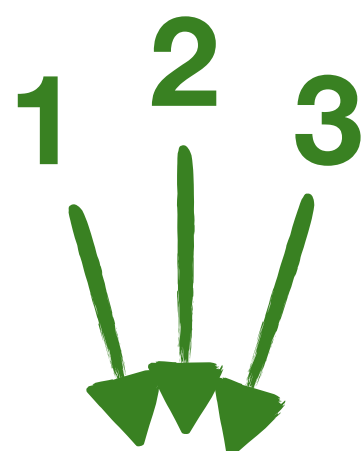


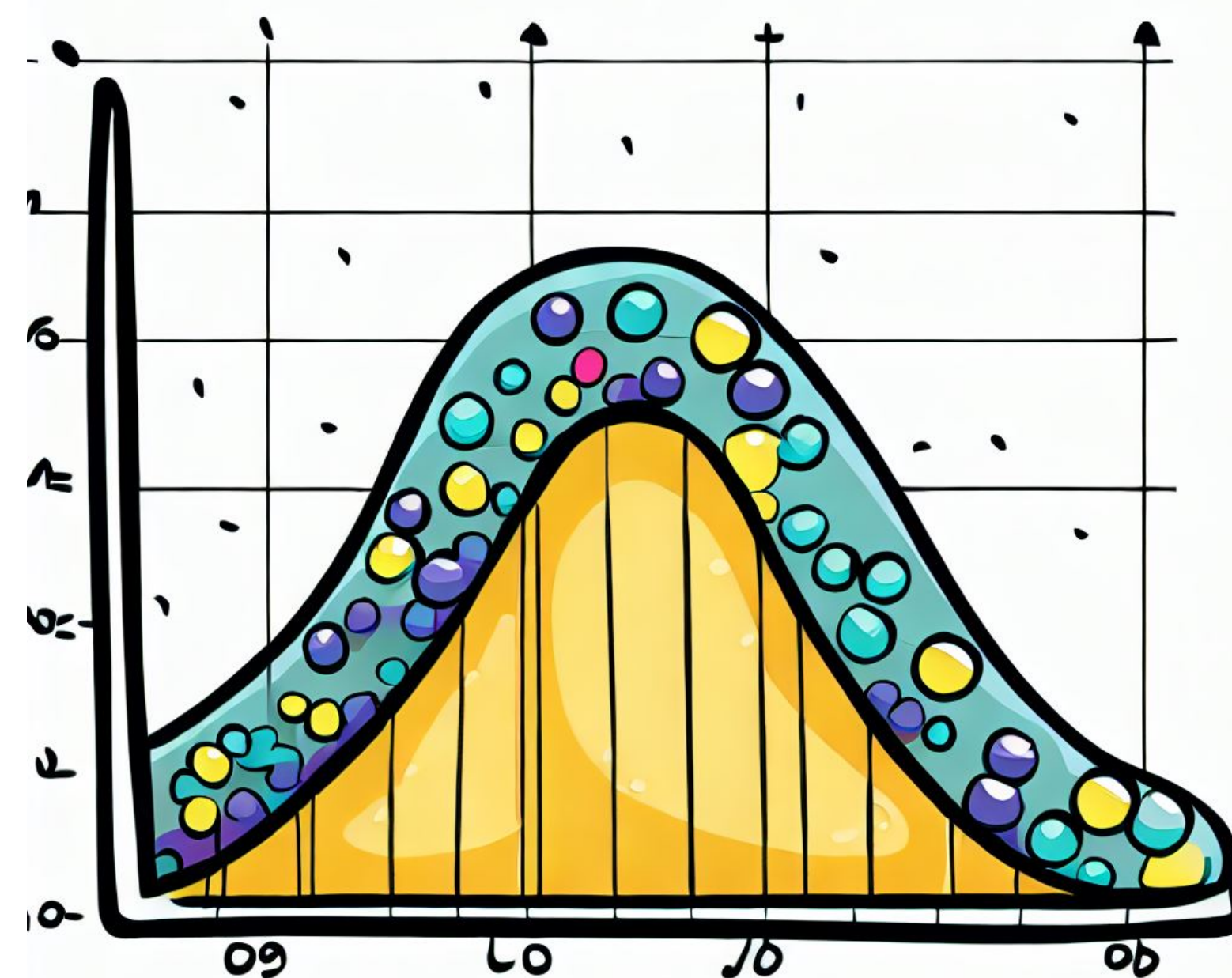
ISTT100y



Uke 6: Normalfordelingen

Thea Bjørnland
Institutt for matematiske fag
NTNU

Husk!
Vi vil ha flere deltakere i referansegruppa
(vi har fra 1002 og 1003)



Microsoft bing prompt:
"Gaussian distribution graph, cartoon"

I dag skal vi...

1. Fullføre sykkel-reperasjonsoppgaven fra uke 5 (normalapproximasjon)
2. Regne på sannsynligheter i normalfordeling, inkludert $aX + bY$ og kritiske verdier
3. Jobbe med tilfeldige utvalg $X_1, X_2, \dots, X_n \rightarrow$ summer og gjennomsnitt

Oppgave 1 (kontinuasjonseksamen 2021)

Anders har nettopp kjøpt seg en brukt el-sykkel. Vi skal studere to typer feil som kan opptre på denne sykkelmodellen; feil med ladekabelen og feil med de elektroniske girene.

Anders antar at feil på ladekabelen kan modelleres som en poissonprosess med rate $\lambda_L = 0.5$ / år, og at feil med girene kan modelleres som en poissonprosess med rate $\lambda_G = 2$ / år. De to prosessene kan antas å være uavhengige.

b) Det koster 3000 kr for å reparere ladekabelen og 1000 kr for å reparere girene hver gang de feiler.

- i. Hvor mye må Anders forvente å betale for reparasjoner av disse i løpet av to år?
- ii. Og hva blir standardavviket i disse reparasjonskostnadene?

Poissonfordeling

$$X \sim \text{Poisson}(\lambda t)$$

$$P(X = x) = \frac{(\lambda t)^x}{x!} e^{-\lambda t},$$

for $x = 0, 1, 2, \dots$

$$E(X) = \lambda t, \quad \text{Var}(X) = \lambda t$$

Oppgave 1 (kontinuasjonseksamen 2021)

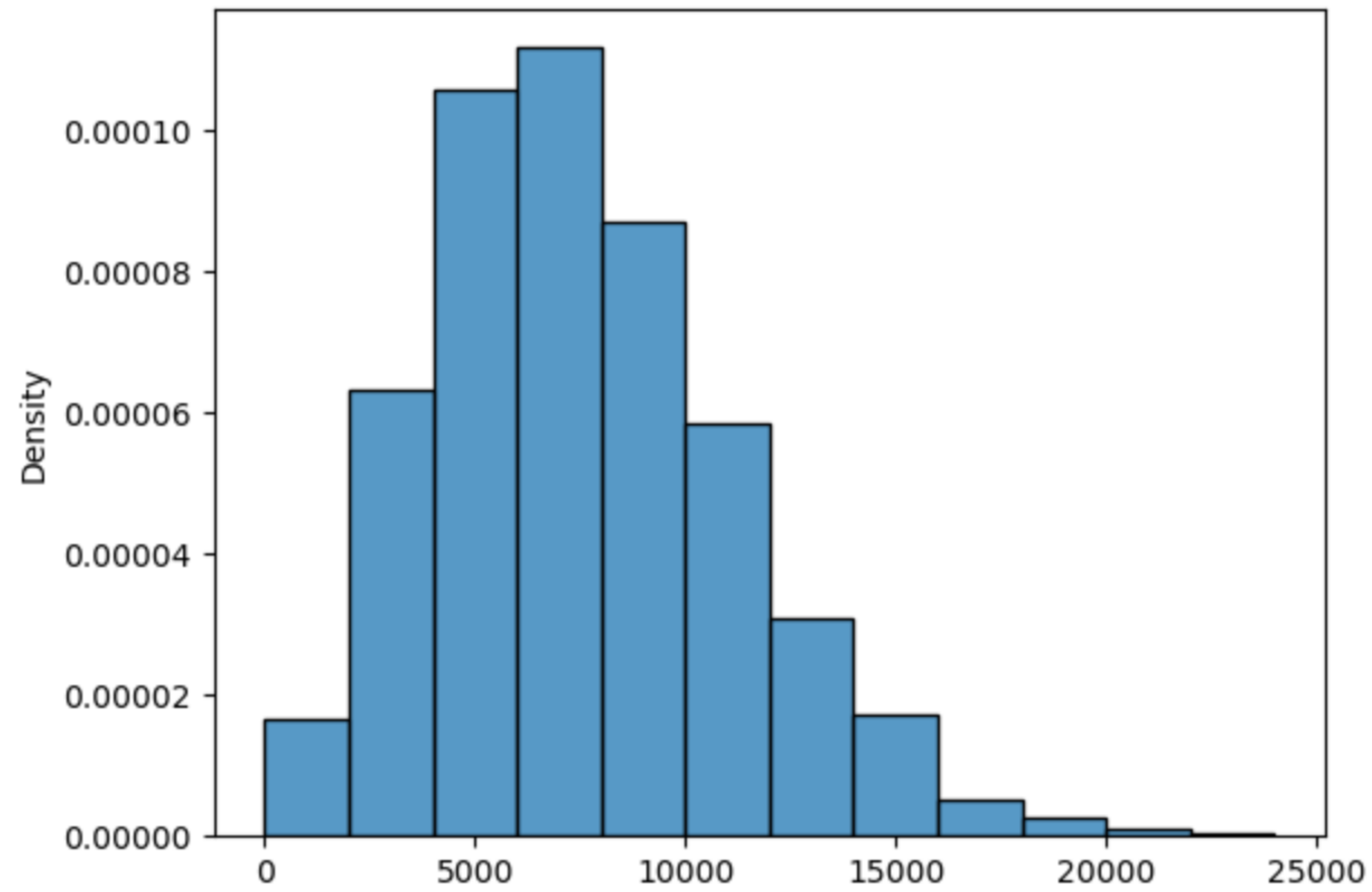
```
import matplotlib.pyplot as plt # plotting
import numpy as np # fordi vi kan trekke fra poissonfordeling
from scipy import stats # statistikk-modulen i scipy-pakken
import seaborn as sns # plotting
```

```
lambda_L = 0.5
lambda_G = 2
t = 2
mu_L = lambda_L*t
mu_G = lambda_G*t
```

```
n = 10000 # Antall simuleringer
y_verdier = []
for _ in range(n):
    x_L = np.random.poisson(mu_L)
    x_G = np.random.poisson(mu_G)
    y = 3000*x_L + 1000*x_G
    y_verdier.append(y)
```

Oppgave 1 (kontinuasjonseksamen 2021)

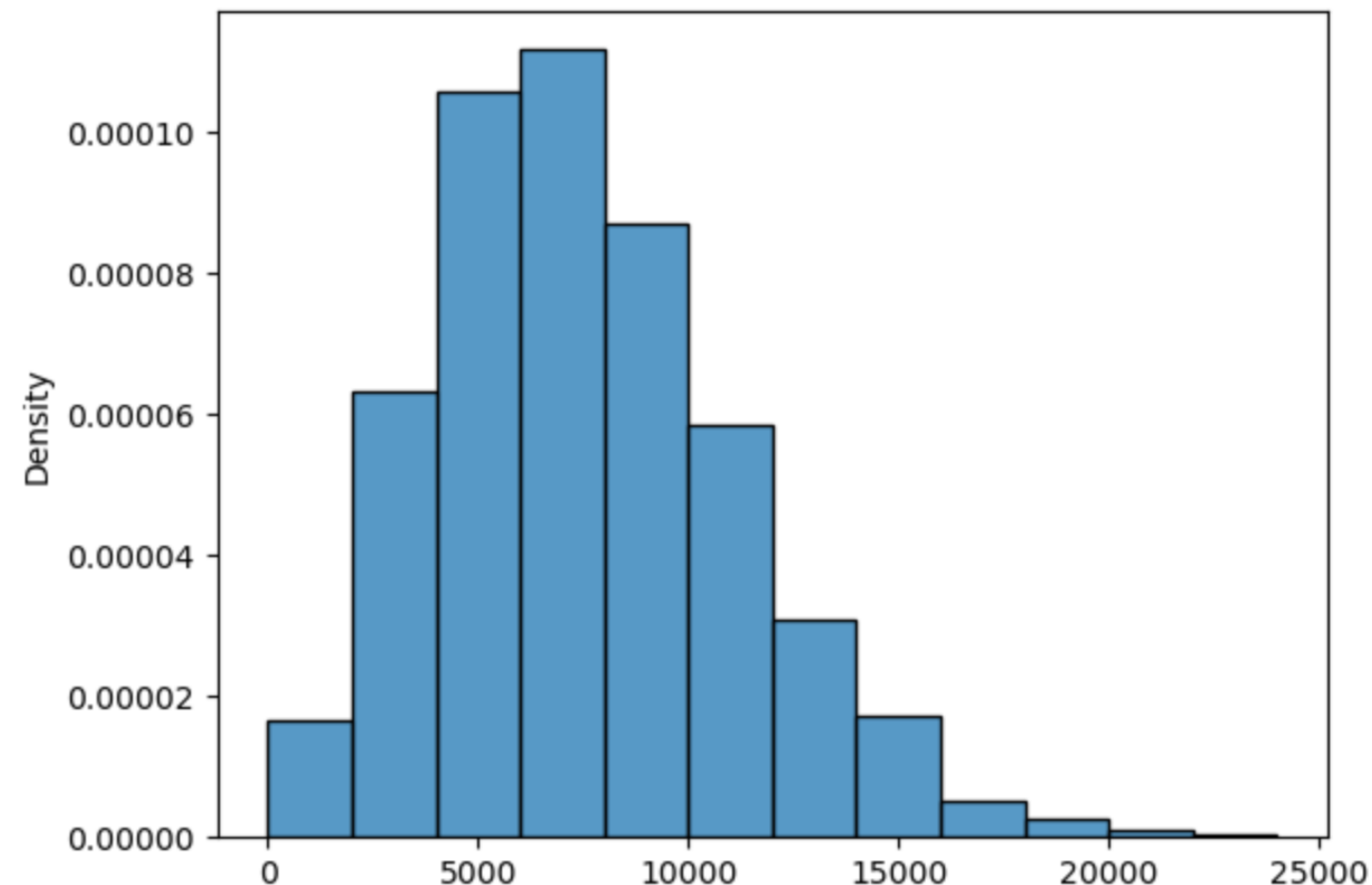
```
sns.histplot(y_verdier, bins = range(0,25000,2000),stat='density')  
plt.show()
```



Oppgave 1(kontinuasjonseksamen 2021)

c) Anders bestemmer seg for at han vil skrote sykkelen dersom kostnadene for reparasjon av ladekabel og gir overskrider 10 000 kr de neste to årene. Bruk *normaltilnærming* for å regne ut sannsynligheten for at Anders skroter sykkelen etter to år.

```
sns.histplot(y_verdier, bins = range(0,25000,2000),stat='density')  
plt.show()
```



Poissonfordeling

$$X \sim \text{Poisson}(\lambda t)$$

$$P(X = x) = \frac{(\lambda t)^x}{x!} e^{-\lambda t},$$

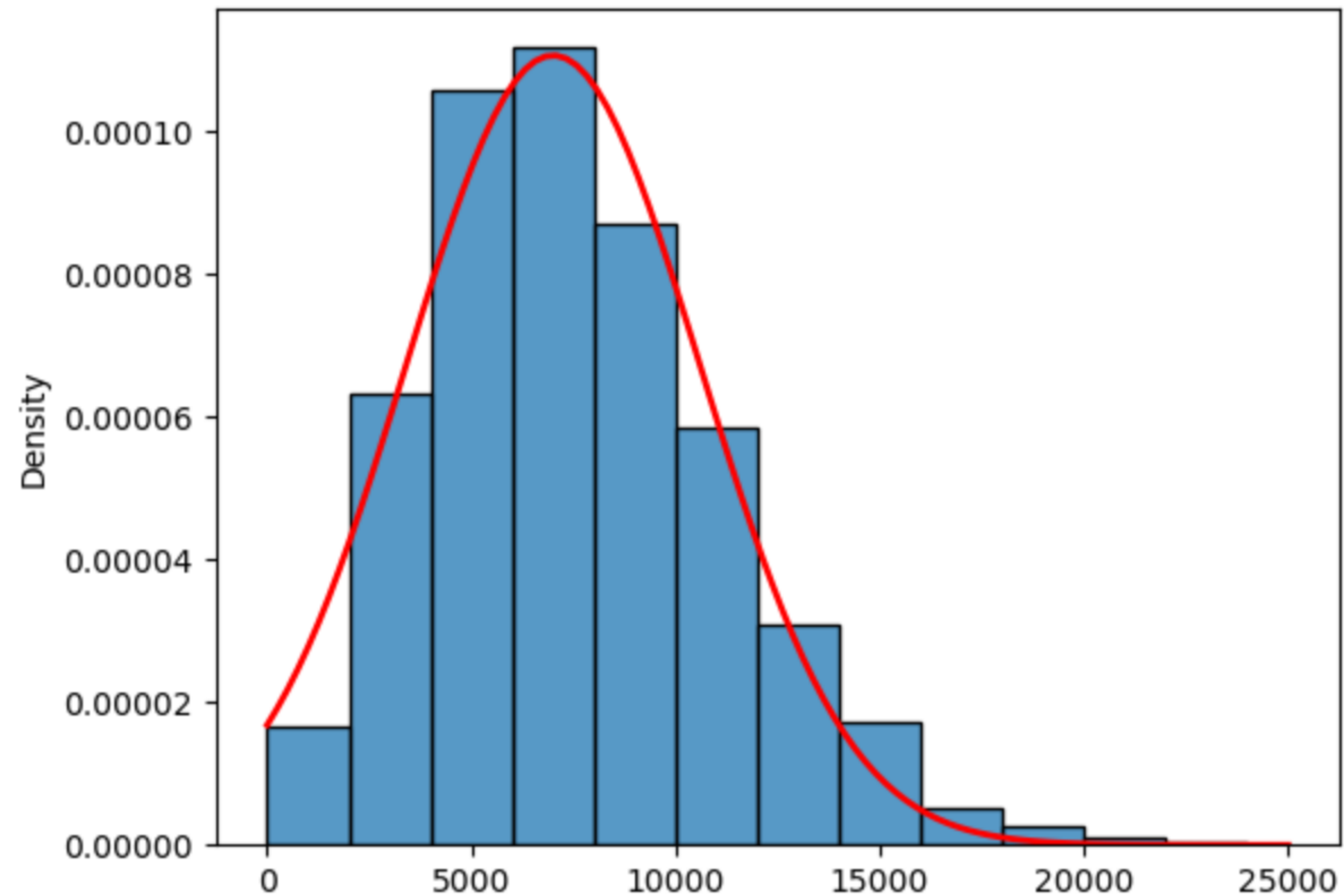
for $x = 0, 1, 2, \dots$

$$E(X) = \lambda t, \quad \text{Var}(X) = \lambda t$$

Oppgave 1(kontinuasjonseksamen 2021)

```
mu = 7000
sigma = 3605.551
yvals = np.linspace(0, 25000, 100)
fy = stats.norm.pdf(yvals, mu, sigma)

sns.histplot(y_verdier, bins = range(0,25000,2000),stat='density')
plt.plot(yvals, fy, 'r-', lw=2, label=f'Normal dist. (mean={mu}, std={sigma})')
plt.show()
```



```
stats.norm.cdf(10000,mu,sigma) #  $P(Y \leq 10000)$ 
```

0.7973097397104328

STANDARD NORMAL DISTRIBUTION: $\Phi(z) = P(Z \leq z)$

z	.._0	.._1	.._2	.._3	.._4	.._5	.._6	.._7	.._8	.._9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389

Oppgave 2

Ved et gartneri dyrkes hodekål. Vekten av den modne kålen, X , antas å være normalfordelt med forventning $\mu = 2$ kg og standardavvik $\sigma = 0.5$ kg

- a) Hvis en velger et kålhode tilfeldig, hva er sannsynligheten for at dette skal veie mellom 2 og 2.5 kg?
- b) Hvis en velger to kålhoder tilfeldig, og deres vekt kan antas uavhengig, hva er sannsynligheten for at differensen i vekt er mindre enn 1 kg?
- c) Hvor mye veier de 5% tyngste kålhodene?
- d) Kålhoder som veier mindre enn 1.5 kg oppfyller ikke kravet til klasse 1 kål. Gitt at et kålhode oppfyller kravet (veier mer enn 1.5 kg), hva er sannsynligheten for at det veier mellom 2 og 2.5 kg?

Kritiske verdier i standard normalfordelingen

$$P(Z > z_{\alpha}) = \alpha$$

α	z_{α}
0.20	0.842
0.10	1.282
0.05	1.645
0.025	1.960
0.02	2.054
0.01	2.326
0.005	2.576
0.001	3.090
0.0005	3.291
0.0001	3.719

Oppgave 3

4.1 Regneregler normalfordeling

Vi ser på n uavhengige stokastiske variabler $X_1, X_2, \dots, X_n$ slik at $X_i \sim N(\mu, \sigma)$, for $i = 1, \dots, n$. Da er

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

og

$$X_1 + X_2 + \dots + X_n \sim N(n\mu, \sqrt{n} \cdot \sigma)$$

a) Vis at når $E(X_1) = E(X_2) = \dots = E(X_n) = \mu$ så er $E(\bar{X}) = \mu$

b) Vis at når $SD(X_1) = SD(X_2) = \dots = SD(X_n) = \sigma$, og X -ene er uavhengige, så er $SD(\bar{X}) = \frac{\sigma}{\sqrt{n}}$

4.1.2 Sentralgrenseteoremet

Dersom $X_1, X_2, \dots, X_n$ er uavhengige stokastiske variabler fra samme sannsynlighetsfordeling med forventning $E(X_i) = \mu$ og varians $\text{Var}(X_i) = \sigma^2$, for $i = 1, \dots, n$, og dersom $n > 30$, så er

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} \text{ tilnærmet } N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

og

$$X_1 + X_2 + \dots + X_n \text{ tilnærmet } N(n\mu, \sqrt{n} \cdot \sigma)$$

Oppgave 4 (eksamen 2023)

En maskin produserer aluminiumsplater som ifølge produktspesifikasjonen skal veie 100 gram. Det blir akseptert et lite avvik fra denne vekten, men dersom vekten avviker mer enn ± 1 gram fra denne verdien, blir aluminiumsplatene vurdert til å være defekt.

Anta at vekten på aluminiumsplatene som produseres er normalfordelt med forventning 100 gram og standardavvik 0.8 gram.

Ti plater pakkes i en pappeske. Pappesken veier 50 gram. Platenes vekt er uavhengige.

a) Hva er forventet vekt av pappesken med ti tilfeldig valgte plater?

b) Hva er standardavviket?

c) I en eske med ti plater, hva er sannsynligheten for at minst 2 plater er defekte?