문제 **1.** 하루에 **30**개의 전구를 제조하는 기계가 있다**.** 해당 기계에서 만들어진 전구의 수명은 **N(100,100)** 분포를 따른다고 가정하자**.**

아래 문제에 대해 파이썬으로 코드를 작성하였다.

1) 하룻동안 생산된 **30**개 전구의 평균 수명을 **R**프로그램을 이용하여 시뮬레이션 해보시오**.**

평균 100, 분산 10이고, Smple Size가 30이 되도록 normal Distribution 생성 그리고 해당 분포의 평균을 산출

2) 동일한 방식으로 **100**일 동안 매일 **30**개 전구의 평균 수명을 측정해 보시오**. 100**일 동안 평균값이 **97 ~ 103** 범위를 벗어난 횟수는 얼마나 되는가**? 100**개의 평균값은 어떤 형태의 분포를 띄는가

위에 구성에서 100번이라는 반복 횟수 추가, 그리고 해당 분포의 평균이 97 ~ 103 밖에
있는 경우 카운트

3) 누군가 각 기계의 유리병 제조 수는 정규분포가 아니라 평균이 **100**인 지수분포**(exponential distribution)**를 따른다고 주장하였다**.** 정규분포 대신 지수분포를 가정하고 **1) ~ 2)**번 과정을 반복하여 실행해 보시오**.**

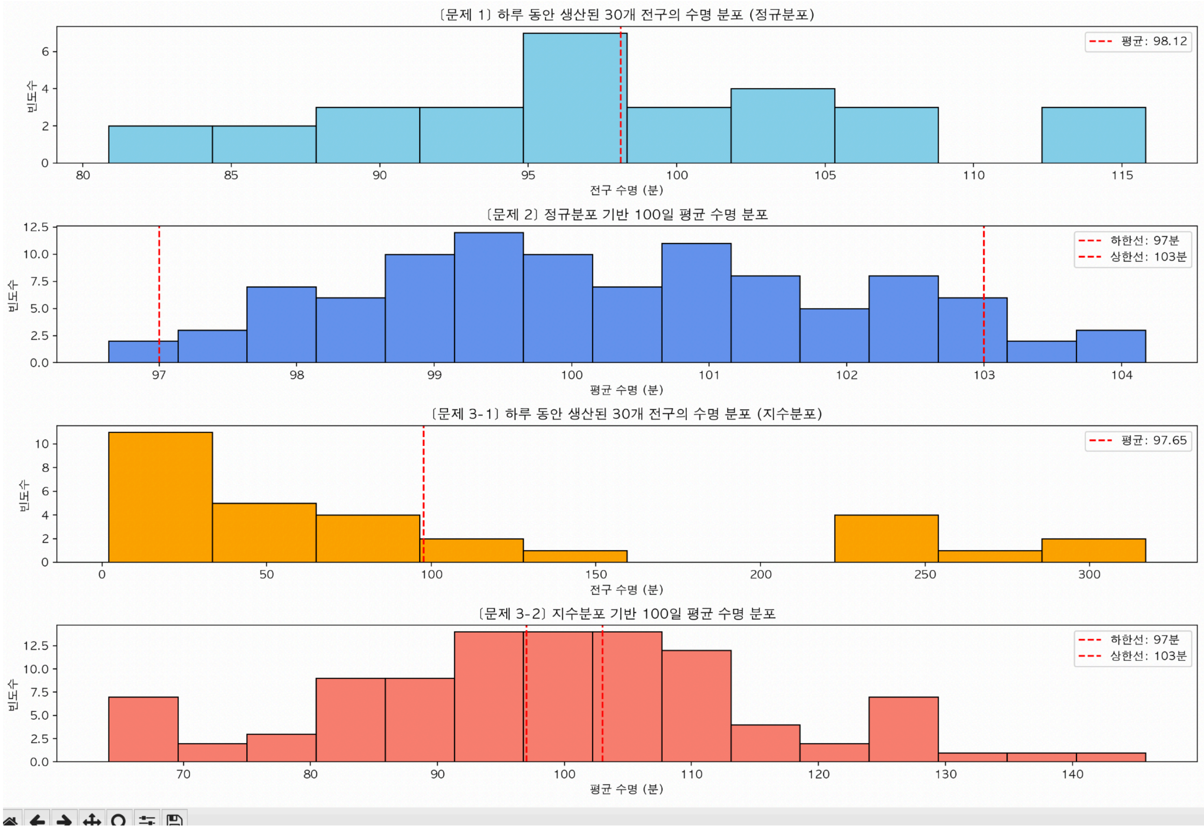
평균이 100인 지수분포 생성 후, 위 과정 수행

```python
import numpy as np
import matplotlib.pyplot as plt
import platform

# 폰트 설정 (Mac용)
if platform.system() == 'Darwin':
   plt.rcParams['font.family'] = 'AppleGothic'
else:
   plt.rcParams['font.family'] = 'Malgun Gothic'

plt.rcParams['axes.unicode_minus'] = False

# 시드 고정
np.random.seed(42)

# 파라미터 설정
mean = 100
std_dev = 10
days = 100
samples_per_day = 30
scale_exp = 100
```

```python
# 문제 [1] - 하루 동안 30개 정규분포 샘플 평균
day1_lifespans = np.random.normal(loc=mean, scale=std_dev, size=samples_per_day)
day1_average = np.mean(day1_lifespans)

# 문제 [2] - 정규분포 기반 100일 평균
lifespan_100days = np.random.normal(loc=mean, scale=std_dev, size=(days, samples_per_day))
daily_means_normal = np.mean(lifespan_100days, axis=1)
outside_range_normal = np.sum((daily_means_normal < 97) | (daily_means_normal > 103))

# 문제 [3] - 지수분포 기반 100일 평균
lifespan_exp_100days = np.random.exponential(scale=scale_exp, size=(days, samples_per_day))
daily_means_exponential = np.mean(lifespan_exp_100days, axis=1)
outside_range_exp = np.sum((daily_means_exponential < 97) | (daily_means_exponential > 103))

# 문제 [3-1] - 하루 동안 30개 지수분포 샘플 평균
day1_lifespans_exp = np.random.exponential(scale=scale_exp, size=samples_per_day)
day1_average_exp = np.mean(day1_lifespans_exp)

# 결과 출력
print("[1] 하루 동안 생산된 30개 전구 평균 수명 (정규분포): {:.2f}분".format(day1_average))
print("[2] 정규분포 기반 100일 평균 중 97~103 범위 벗어난 날 수: {}일".format(outside_range_normal))
print("[3-1] 하루 동안 생산된 30개 전구 평균 수명 (지수분포): {:.2f}분".format(day1_average_exp))
print("[3-2] 지수분포 기반 100일 평균 중 97~103 범위 벗어난 날 수: {}일".format(outside_range_exp))


# 문제별 그래프 그리기
plt.figure(figsize=(15, 10))

# 문제 1: 하루 평균 수명 (정규분포)
plt.subplot(4, 1, 1)
plt.hist(day1_lifespans, bins=10, color='skyblue', edgecolor='black')
plt.axvline(day1_average, color='red', linestyle='--', label=f"평균: {day1_average:.2f}")
plt.title("[문제 1] 하루 동안 생산된 30개 전구의 수명 분포 (정규분포)")
plt.xlabel("전구 수명 (분)")
plt.ylabel("빈도수")
```

```python
plt.legend()

# 문제 2: 정규분포 기반 100일 평균
plt.subplot(4, 1, 2)
plt.hist(daily_means_normal, bins=15, color='cornflowerblue', edgecolor='black')
plt.axvline(97, color='red', linestyle='--', label="하한선: 97분")
plt.axvline(103, color='red', linestyle='--', label="상한선: 103분")
plt.title("[문제 2] 정규분포 기반 100일 평균 수명 분포")
plt.xlabel("평균 수명 (분)")
plt.ylabel("빈도수")
plt.legend()

# 문제 3-1: 하루 평균 수명 (지수분포)
plt.subplot(4, 1, 3)
plt.hist(day1_lifespans_exp, bins=10, color='orange', edgecolor='black')
plt.axvline(day1_average_exp, color='red', linestyle='--', label=f"평균:
{day1_average_exp:.2f}")
plt.title("[문제 3-1] 하루 동안 생산된 30개 전구의 수명 분포 (지수분포)")
plt.xlabel("전구 수명 (분)")
plt.ylabel("빈도수")
plt.legend()

# 문제 3-2: 지수분포 기반 100일 평균
plt.subplot(4, 1, 4)
plt.hist(daily_means_exponential, bins=15, color='salmon', edgecolor='black')
plt.axvline(97, color='red', linestyle='--', label="하한선: 97분")
plt.axvline(103, color='red', linestyle='--', label="상한선: 103분")
plt.title("[문제 3-2] 지수분포 기반 100일 평균 수명 분포")
plt.xlabel("평균 수명 (분)")
plt.ylabel("빈도수")
plt.legend()



plt.tight_layout()
plt.show()
```

```
(6.86x) (base) kimdawoon@gimdaun-ui-MacBook-Pro ~/DataScience_Statistic_R_Data/HW/W06 main python problem1.py
[1] 하루 동안 생산된 30개 전구 평균 수명 (정규분포): 98.12분
[2] 정규분포 기반 100일 평균 중 97~103 범위 벗어난 날 수: 9일
[3-1] 하루 동안 생산된 30개 전구 평균 수명 (지수분포): 97.65분
[3-2] 지수분포 기반 100일 평균 중 97~103 범위 벗어난 날 수: 85일
```



문제 **2.** 하루에 **12**시간 운영하는 공장에서 제품을 생산하고 있다**.** 제품 생산을 위한 장비의 필수 부품 중에 매일 교체해 줘야 하는 소모품 한 개가 있는데**,** 해당 부품의 수명은 지수분포**(exponential distribution)**을 따르고**,** 평균 수명은 **15**시간이라고 한다**.** 해당 소모품이 공장 운영 시간 **(12**시간**)** 중에 고장나 운영에 차질이 발생하지 않도록 하기 위해 매일 추가로 하나의 소모품을 더 준비해 놓고 문제가 발생하면 바로 교체할 수 있다고 하자**.**

1) 공장 운영 시간 **(12**시간**)** 이내에 여분으로 준비해 놓은 소모품까지 모두 고장나서 생산에 차질이 발생할 확률은 얼마인가**?**

$\lambda = 1/15$

여분 1개까지 포함하여 2개가 있다.
12시간 내에 2개가 고장날 확률을 계산하면 된다.

$R(t) = P(T > t) = e^{-\lambda t}$

$\lambda = 1/15,\ t = 12$

$R(12) = e^{-12/15} \approx 0.4493$

44.93%는 한 개의 소모품이 12시간 내에 고장나지 않을 확률이다.

제품이 총 2개이므로, 또한 지수분포는 무기억성 특징을 가지고 고장을
독립으로 보면
아래 와 같이 계산할 수 있다.

$$1 - R(12))^2 \approx (0.5507)^2 \approx 0.3032$$

즉, 30.32%이다.

2) 생산 차질 확률을 **1%** 이하로 낮추기 위해서는 소모품 여분을 몇 개 준비해야 하는가**?**

지수분포이므로, 사건은 독립이다.
따라서 만약에 n+1(여기서 n은 소모품 갯수)의 제품이 있을 때, 12시간 내에
전부 고장날
확률은 다음과 같다.

$$P_{전체고장} = (P_{개별고장})^{n+1}$$

여기서 고장이 0.01 이하가 되기 위한 n을 구하기 위해서는

$$(P_{개별고장})^{n+1} \le 0.01$$

양변에 로그를 취하면,

$$(n + 1) \cdot log(P_{개별고장}) \le log(0.01)$$
$$n + 1 \ \ge \frac{log(0.01)}{log(P_{개별고장})} \approx \frac{-2}{log(0.5507)} \approx 8$$

따라서 n은 7이다.