DAMA50 - 5th assignment

Quiz 1

```
#1
a = 0.85 - 0.27 - 0.19
b = 1 - (0.85 - 0.27 - 0.19) - 0.03

print(a)
print(b)
```

```
0.3900000000000000
0.5800000000000000
```

Quiz 2

```
#2
from scipy.special import comb

# Define the given values
n = 6
p = 0.2
k = 4

# Calculate the probability of getting more than half correct by chance
probability = sum(comb(n, i, exact=True) * (p**i) * ((1 - p)**(n - i)) for i in range(k + 1, n + 1))

probability
```

```
0.0016000000000000000
```

Quiz 4

```
#4
P_A = 0.30
P_B = 0.40
P_B_given_A = 0.20

P_A_and_B = P_A * P_B_given_A

P_A_and_B_percent = P_A_and_B * 100

print(P_A_and_B_percent)
```

```
6.000000000000000
```

DAMA50 - 5th assignment

Quiz 5

```
#5
EX = 109/50
EY = 157/100
EXY = 171/50

cov_XY = EXY - (EX * EY)
#we see that data are consistent so it's not c
if cov_XY == 0:
    answer = "X & Y uncorrelated (a)"
elif cov_XY > 0:
    answer = "X & Y positively correlated (e)"
elif cov_XY < 0:
    answer = "X & Y negatively correlated (d)"
else:
    answer = "Inconsistent data (b)"

answer
```

```
'X & Y negatively correlated (d)'
```

Quiz 6

```
#6
var_X = 10
var_Y = 2

var_X_minus_2Y_plus_1 = 1 + (-2)**2 * var_Y + var_X

var_X_minus_2Y_plus_1
```

```
19
```

DAMA50 - 5th assignment

Problem 7

0.015 seconds Explain Format Copy 1

In [42]:
```
from scipy.stats import norm

m = 70
s = 14
min_75 = 75
max_90 = 90

z_min = (min_75 - m) / s
z_max = (max_90 - m) / s

perc = (norm.cdf(z_max) - norm.cdf(z_min)) * 100

perc
```

Out[42]:
```
28.392870544100056
```

0.014 seconds Explain Format Copy 2

In [43]:
```
from scipy.stats import norm

prob_7 = 0.06
prob_22 = 0.75

z_7 = norm.ppf(prob_7)

z_22 = norm.ppf(prob_22)

m = (7 - z_7 * 7) / (1 - z_7)
s = (22 - z_22 * 22) / (1 - z_22)

m, s
```

Out[43]:
```
(7.0, 22.000000000000004)
```

0.007 seconds Explain Format Copy 3

In [44]:
```
prob_7_verified = norm.cdf(7, loc=m, scale=s)
prob_22_verified = norm.cdf(22, loc=m, scale=s)

prob_7_verified, prob_22_verified
```

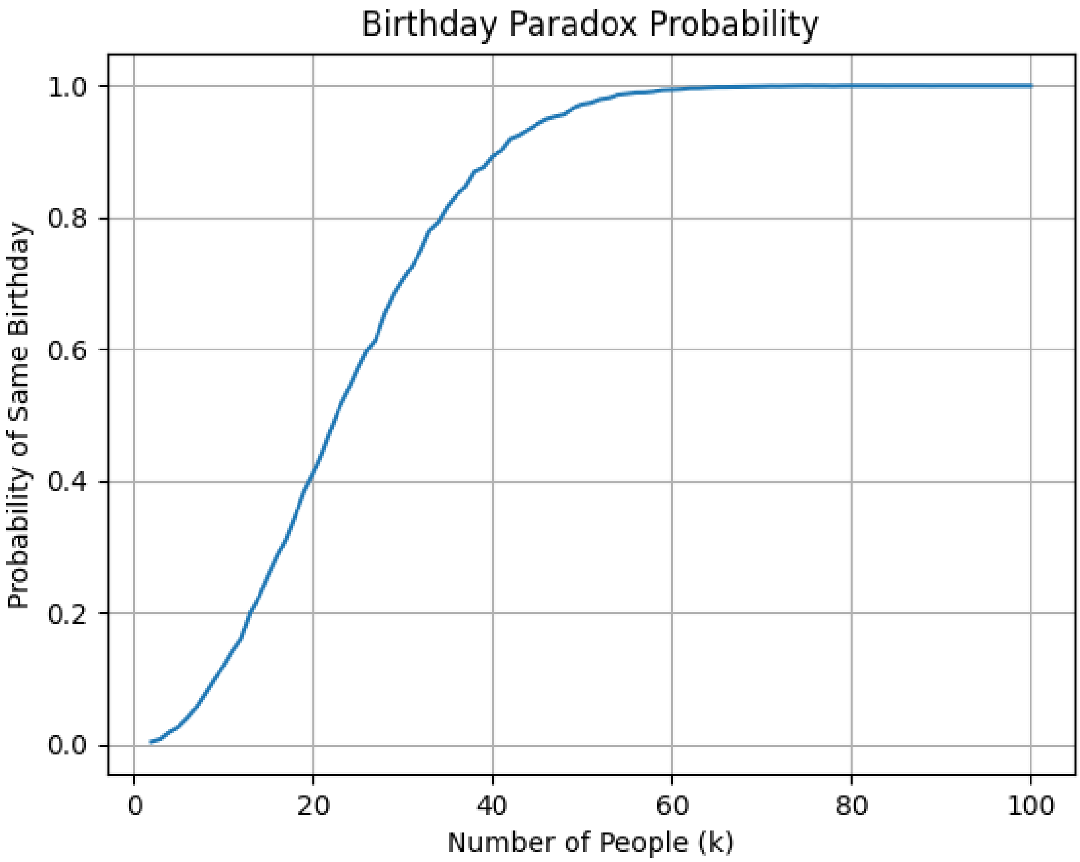
Out[44]:
```
(0.5, 0.7523230370047824)
```

DAMA50 - 5th assignment

Problem 8

```python
import random
import matplotlib.pyplot as plt

def sagebdate(k):
    birthdays = set()
    for i in range(k):
        birthday = random.randint(1, 365)
        if birthday in birthdays:
            return 1
        birthdays.add(birthday)
    return 0

def p(k, n):
    count = sum(sagebdate(k) for _ in range(n))
    return count / n

n = 10000
ks = list(range(2, 101))
probs = [p(k, n) for k in ks]

plt.plot(ks, probs)
plt.xlabel('Number of People (k)')
plt.ylabel('Probability of Same Birthday')
plt.title('Birthday Paradox Probability')
plt.grid(True)
plt.show()
```

DAMA50 - 5th assignment



DAMA50 - 5th assignment

Problem 9

All numbers are rounded to 4 decimals

9) α) $P(x=0) = 1211/102354 = 0{,}0119$

$P(x=1) = 33279/102354 = 0{,}3249$

$P(x=2) = 38862/102354 = 0{,}3798$

$P(x=3) = 18837/102354 = 0{,}1838$

$P(x=4) = 7684/102354 = 0{,}0752$

$P(x=5) = 2482/102354 = 0{,}0244$

β) $F(x=0) = P(x \leq 0) = P(x=0) = 0{,}0119$

$F(x=1) = P(x \leq 1) = P(x=0) + P(x=1) = 0{,}0119 + 0{,}3249 =$
$= 0{,}3368$

$F(x=2) = P(x \leq 2) = P(x=0) + P(x=1) + P(x=2)$
$= 0{,}0119 + 0{,}3249 + 0{,}3798 = 0{,}7166$

$F(x=3) = P(x \leq 3) = P(x=0) + P(x=1) + P(x=2) + P(x=3)$
$= 0{,}0119 + 0{,}3249 + 0{,}3798 + 0{,}1838$
$= 0{,}9004$

$F(x=4) = P(x \leq 4) = P(x=0) + P(x=1) + P(x=2) + P(x=3) + P(x=4)$
$= 0{,}0119 + 0{,}3249 + 0{,}3798 + 0{,}1838 + 0{,}0752$
$= 0{,}9756$

$F(x=5) = P(x \leq 5) = P(x=0) + P(x=1) + P(x=2) + P(x=3) + P(x=4) + P(x=5)$
$= 0{,}0119 + 0{,}3249 + 0{,}3798 + 0{,}1838 + 0{,}0752 + 0{,}0244$
$= 1$

c) $P(x=4 \mid x \geq 2) = \dfrac{P(x=4 \text{ and } x \geq 2)}{P(x \geq 2)}$

- $P(x=4 \text{ and } x \geq 2) = P(x=4) = 0{,}0752$

- $P(x \geq 2) = P(x=2) + P(x=3) + P(x=4) + P(x=5) =$
$= 0{,}3798 + 0{,}2838 + 0{,}0752 + 0{,}0244$
$= 0{,}6634$

So $P(x=4 \mid x \geq 2) = 0{,}0752 / 0{,}6634 = 0{,}1133$

d) $E(x) = \sum_{x=0}^{5} x \cdot P(x) = 0 \cdot P(x=0) + 1 \cdot P(x=1) + 2 \cdot P(x=2) + 3 \cdot P(x=3) + 4 \cdot P(x=4)$
$+ 5 \cdot P(x=5)$
$= 0 \cdot 0{,}0119 + 1 \cdot 0{,}3249 + 2 \cdot 0{,}3798 + 3 \cdot 0{,}1838 + 4 \cdot 0{,}0752$
$+ 5 \cdot 0{,}0244$.
$= 0 + 0{,}3249 + 0{,}7596 + 0{,}5514 + 0{,}3008 + 0{,}1220$
$= 2.0587$

e) $Var(x) = \sum_{x=0}^{5} (x-\mu)^2 \cdot P(x)$ where $\mu = E(x)$

$= \big((0-2.0587)^2 \cdot 1212 + (1-2.0587)^2 \cdot 33279 + (2-2.0587)^2 \cdot$
$\cdot 38861 + (3-2.0587)^2 \cdot 18837 + (4-2.0587)^2 \cdot 7684 +$
$+ (5-2.0587)^2 \cdot 2482\big) / 102354$

$= \dfrac{5.132{,}5255 + 37.300{,}6237 + 133{,}9029 + 16690{,}4426}{102354}$

$$+\frac{28{,}958.2735 + 21{,}472{,}3928}{102354} = \frac{109\,688.15}{102354}$$

$$= 1.0716$$

DAMA50 - 5th assignment

Problem 10

10)a) $E(x) = \int_{4}^{10} x \cdot f(x) \cdot dx$

$= \int_{4}^{10} x \cdot \left(1/72 \cdot (-x^2+12x-20\right) dx$

$= \int_{4}^{10} 1/72 (-x^3+12x^2-20x) \; dx$

$= \frac{1}{72} \int_{4}^{10} (-x^3+12x^2-20x) \; dx$

$= \frac{1}{72} \cdot \left(\frac{-x^4}{4} + 4x^3 - 10x^2 \right)$ from $x=4$ to $x=10$

$= \frac{1}{72} \left(\left(-\frac{10^4}{4} + 4 \cdot 10^3 - 10 \cdot 10^2 \right) - \left(-\frac{4^4}{4} + 4 \cdot 4^3 - 10 \cdot 4^2 \right) \right)$

$= \frac{1}{72} \left(-2500 + 4000 - 1000) - (-256 + 256 - 160) \right)$

$= \frac{1}{72} (500 + 160) = \frac{660}{72} = 9.16\bar{6}$

b) $Var(x) = E(x^2) - (E(x))^2$

$E(x^2) = \int_{4}^{10} x^2 f(x) dx$

$= \int_{4}^{10} x^2 \cdot \frac{1}{72} (-x^4 + 12x^3 - 20x^2) dx$

$= \frac{1}{72} \left(-x^5/5 + 3x^4 - \frac{20x^3}{3} \right)$ from $x=4$ to $x=$

$= \frac{1}{72} \left(\left(-\frac{10^5}{5} + 3 \cdot 10^4 - \frac{20 \cdot 10^3}{3} \right) - \left(-\frac{4^5}{5} + 3 \cdot 4^4 - \frac{20 \cdot 4^3}{3} \right) \right)$

$$= \frac{1}{72}\left((-20000 + 30000 - 6666.67) - (-204.8 + 768 - 426.66)\right)$$

$$= \frac{1}{72}(3333.3333 - 136.5333)$$

$$= \frac{1}{72}(3296.8000)$$

$$= 44{,}4$$

c) $\sigma = \sqrt{Var(x)} = \sqrt{44{,}4} = 6{,}663$

$$P(\mu - \sigma < X < \mu) = P(9{,}1667 - 6{,}663 < X < 9{,}1667)$$
$$= P(2.5037 < X < 9{,}1667)$$

$$F(x) = \int_4^x f(t)\,dt$$

$$= \frac{1}{72}\int_4^x (-t^2 + 12t + 20)\,dt$$

$$= \frac{1}{72}\left(-\frac{t^3}{3} + 6t^2 - 20t\right) \text{ from } t=4 \text{ to } t=x$$

$$= \frac{1}{72}\left(\left(-\frac{x^3}{3} + 6x^2 - 20x\right) - \left(-\frac{4^3}{3} + 6\cdot 4^2 - 20\cdot 4\right)\right)$$

$$= \frac{1}{72}\left(-\frac{x^3}{3} + 6x^2 - 20x + 5.3\bar{3}\right)$$

$$P(\mu - \sigma < x < \mu) = P(2{,}5037 < x < 9{,}1667)$$
$$= F(9{,}1667) - F(2{,}5037) = \ldots$$