

**Definition:** Let  $X$  be the number of letter's addresses that the machine fails to correctly read.

**Distribution:**  $X \sim \text{Binomial}(n=4000, p=0.0021)$

**Probability:**  $P(X \geq 10)$

- a) **Justification:** The Poisson approximation is appropriate when  $n$  is large, and  $p$  is small such that  $\lambda=np$  is moderate. Here,  $n=4000$  and  $p=0.0021$ , giving  $\lambda=4000 \times 0.0021=8$ .

**Probability Calculation:**  $P(X \geq 10) \approx 1 - P(X \leq 9)$

```
lambda <- 4000 * 0.0021
prob <- 1 - ppois(9, lambda=lambda)
prob
[1] 0.3340803
```

**Poisson Approximation:**  $P(X \geq 10) \approx 0.3340803$

- b) The normal approximation is appropriate when  $np > 5$  and  $n(1-p) > 5$ .

Checking:  $np = 4000 \times 0.0021 = 8.4$  (which is  $> 5$ )

$n(1-p) = 4000 \times 0.9979 = 3991.6$  (which is  $> 5$ )

Therefore, the normal approximation is appropriate.

Mean ( $\mu$ ) =  $np = 8.4$

$$\sqrt{np(1-p)} = \sqrt{(4000 \times 0.0021) \times (1 - 0.0021)}$$

Standard deviation ( $\sigma$ ) =  $\sqrt{np(1-p)} = \sqrt{8.4 \times 0.9979} \approx 2.895230$

With continuity correction, we calculate  $P(X \geq 9.5)$  instead of  $P(X \geq 10)$ :

```
> 1 - pnorm(9.5, mean = 8.4, sd = 2.895230)
```

```
> [1] 0.3519967
```

**Normal Approximation with Continuity Correction:**  $P(X \geq 9.5) \approx 0.3519967$

- c) Using the same mean and standard deviation as in (b), but without continuity correction:

```
> 1 - pnorm(10, mean = 8.4, sd = 2.895230)
```

```
> [1] 0.2902573
```

**Normal Approximation without Continuity Correction:**  $P(X \geq 10) \approx 0.2902573$

- d) Calculating the true probability using R's binomial CDF:

```
> 1 - pbinom(9, size = 4000, prob = 0.0021)
```

```
> [1] 0.3339988
```

**True Probability (Binomial CDF):**  $P(X \geq 10) = 0.3339988$

Therefore, by comparison of the values from a, b and c that the value from (a) or Poisson approximation (0.3340803) was the closest to the true value (d) (0.3339988).