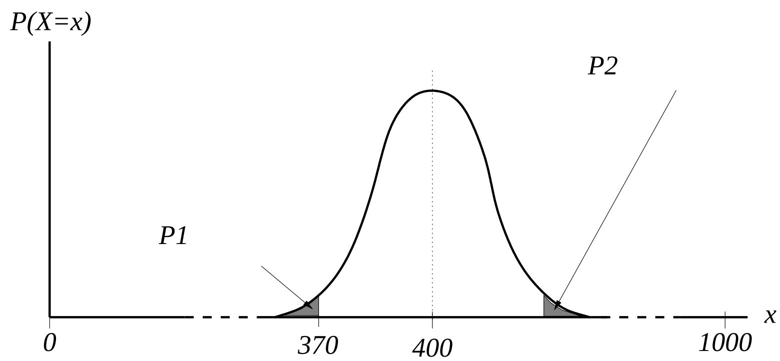
1.(a) Let X be the number of Ardent-voters out of 1000 respondents.

$$X \sim \text{Binomial}(1000,\ p_A)$$
$$H_0 : p_A = 0.4.$$
$$H_1 : p_A \neq 0.4. \qquad \text{(two-sided test)}$$

(b, c) Under H_0, $X \sim \text{Binomial}(1000, 0.4)$.
The probability function of X peaks at about $1000 \times 0.4 = 400$.



(d) For the p-value, $\mathbb{P}(X \leq 370) = F_X(370)$.
Thus the R command for the p-value is: `2 * pbinom(370, 1000, 0.4)`

(e) Interpretation: about 5.6% of the time when the true value of p is 0.4, the polling of 1000 people will give an answer as extreme as $x = 370$. This p-value is small, but not extremely small. Therefore:
(i) We have some evidence against the null hypothesis that $p = 0.4$.
(ii) We do *not* have proof against H_0. (We never get *definite* proof by this method, but this p-value of 0.056 does not come anywhere close to proof.)
(iii) The observed polling *is* compatible with H_0. A result as extreme as 370/1000 can be expected 5.6% of the time under H_0, so it is certainly compatible with H_0. It is *quite unusual* under H_0, but still *compatible* with H_0.

(f) From the results, we see that both Labour *and* National's results could be compatible with them having equal support at 40%. Clearly they could also be compatible if Labour were a little ahead of National (e.g. 40% for Labour, 39% for National).

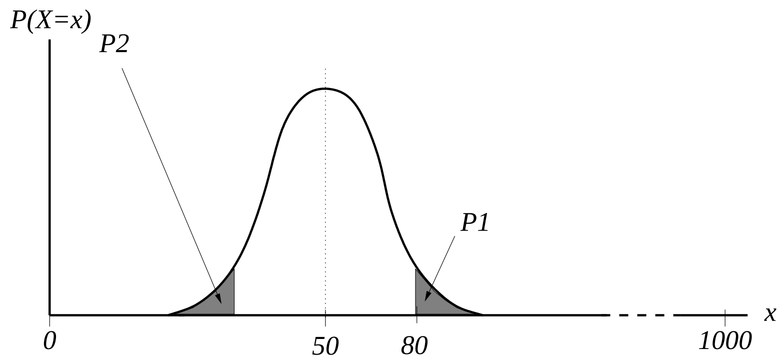
Note that this refers to just one poll: if repeated polls continue to put National ahead of Labour, we have good reason to believe that National really is ahead of Labour in the population.

2.(a) Let X be the number of Winsome-voters out of the 1000 respondents.

$$\begin{aligned} X &\sim \text{Binomial}(1000,\ p) \qquad \text{where } p \text{ is unknown.} \\ H_0 &: \ p = 0.05. \\ H_1 &: \ p \neq 0.05. \qquad \text{(two-sided test)} \end{aligned}$$

(b, c) Under H_0, $X \sim \text{Binomial}(1000, 0.05)$.

The probability function of X peaks at about $1000 \times 0.05 = 50$.



(d) For the p-value,

$$\begin{aligned} \mathbb{P}(X \geq 80) &= 1 - \mathbb{P}(X < 80) \\ &= 1 - \mathbb{P}(X \leq 79) \\ &= 1 - F_X(79). \end{aligned}$$

Thus the R command for the p-value is: `2 * (1 - pbinom(79, 1000, 0.05))`

(e) Interpretation: when the true value of p is 0.05, it is **extremely** unlikely for a polling of 1000 people to give an answer as extreme as $x = 80$. Therefore we have extremely strong evidence against the null hypothesis that $p = 0.05$.

The observed polling is **barely compatible** with the hypothesis that the true level of support is 5%. *(Answers saying 'not compatible' will be accepted for this question.)*

(f) The conclusions are very different for the two cases. The difference arises because the hypothesized value of the Binomial probability p is very small (0.05) in the second case, and much larger (0.4) in the first case.

Explanation: *not needed to get the marks for this question.* When the Binomial probability p is close to 0 or 1, the Binomial distribution has low variance, so it cannot tolerate large deviations away from the population proportion. This explains why a sample deviation of 30/1000 responses away from the hypothesized population p gives no evidence against $H_0 : p = 0.4$, but extremely strong evidence against $H_0 : p = 0.05$.

3.(a) $X \sim \text{Binomial}(7, p)$.

(b) $x = 5$.

(c) $L(p\,;5) = \mathbb{P}(X = 5)$ when $X \sim \text{Binomial}(7, p)$.

Thus

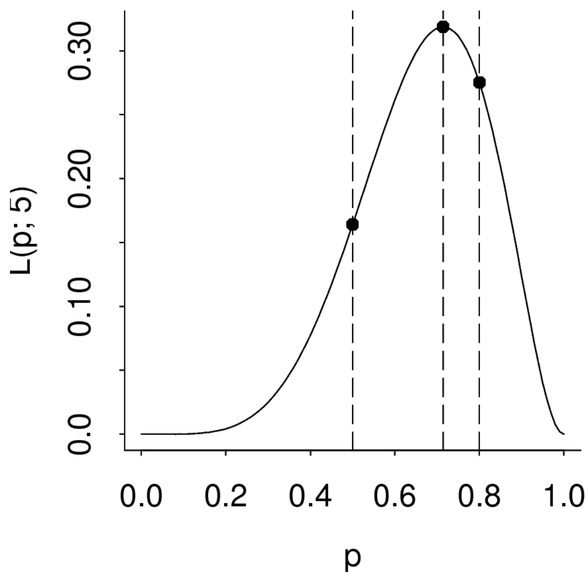
$$\begin{aligned} L(p\,;5) &= \binom{7}{5} p^5 (1-p)^{7-5} \\ &= 21p^5(1-p)^2 \quad \text{for } 0 < p < 1. \end{aligned}$$

(d) $L\left(\frac{1}{2}\,;5\right) = 21 \times \left(\frac{1}{2}\right)^5 \times \left(\frac{1}{2}\right)^2 = 0.164$.

(e) $L\left(\frac{5}{7}\,;5\right) = 21 \times \left(\frac{5}{7}\right)^5 \times \left(\frac{2}{7}\right)^2 = 0.319$.

(f) $L\,(0.8\,;5) = 21 \times (0.8)^5 \times (0.2)^2 = 0.275$.

(g)



(h)

$$\begin{aligned} \frac{dL}{dp} &= 21 \times \Big\{ 5p^4(1-p)^2 + p^5 \times 2(1-p) \times (-1) \Big\} \\ &= 21p^4 \Big\{ 5(1-p)^2 - 2p(1-p) \Big\} \\ &= 21p^4(1-p) \Big\{ 5(1-p) - 2p \Big\} \\ &= 21p^4(1-p) \Big\{ 5 - 7p \Big\} \qquad \text{as stated.} \end{aligned}$$

(i) The maximizing value of p satisfies:

$$\begin{aligned} \frac{dL}{dp} &= 0 \\ \Rightarrow \quad 21p^4(1-p)(5-7p) &= 0 \\ \Rightarrow \quad p = 0, \quad p = 1, \quad \text{or} \quad p &= \frac{5}{7}. \end{aligned}$$

We can see from the sketch that $p \neq 0$ and $p \neq 1$, so the maximizing value of p is

$$p = \frac{5}{7}.$$

(j) The maximum likelihood estimate, $p = \frac{5}{7}$, is the value of p at which the observation $X = 5$ is more likely than at any other value of p.

4.(a) $L(p\,; 4) = \mathbb{P}(X = 4)$ when the parameter takes the value p.
So

$$L(p\,; 4) = (1-p)^3 p \qquad \text{for } 0 < p < 1.$$

(b)

$$\begin{aligned} \frac{dL}{dp} &= -3(1-p)^2 p + (1-p)^3 \\ &= (1-p)^2 \left\{ -3p + (1-p) \right\} \\ &= (1-p)^2(1-4p) \qquad \text{as stated.} \end{aligned}$$

(c) The maximizing value of p satisfies:

$$\begin{aligned} \frac{dL}{dp} &= 0 \\ \Rightarrow \quad (1-p)^2(1-4p) &= 0 \\ \Rightarrow \quad p = 1, \quad &\text{or} \quad p = \frac{1}{4}. \end{aligned}$$

We can see from the sketch that $p \neq 1$, so the maximizing value of p is

$$p = \frac{1}{4}.$$

(d) Sammie first succeeds on his 4th jump, so his probability of success could be estimated as 1/4 for each jump. This is the same as the maximum likelihood estimate of 1/4.
