

b).i)

Using the nle() command in SGC and counting the number of ^{1's} I estimate that the number of packs with ^{at least} one colour missing is $\approx \frac{6}{100} = 6\%$ ✓

c) $\frac{4}{4}$

³ From the graph, the triangular model can be a very ^{at best} approximate fit (The first two bars do not fit the triangular model). ^{Don't be afraid to be quite critical of a model}
ii) Try a value for θ of 110 (the largest value on the horizontal axis). ^{Note also the 3 exceptionally large values between 100-110}

$$P(W > 80) = 1 - F(80) = 1 - \frac{\binom{80}{110}}{\binom{121}{121}} = 1 - \frac{64}{121} = 57$$

³ The above result disagrees with a visual estimate obtained from the graph, which would be much smaller, suggesting a smaller parameter for θ , of around 85 million words. This also casts doubt on the suitability of the triangular model.