Paired data

Gather 100 rats and determine whether they are infected with viruses A and B.

	I-B	NI-B	
I-A	9	9	18
NI-A	20	62	82
	29	71	100

Underlying probabilities

		B: 0	B: 1	
A	0	p_{00}	p_{01}	p_{0+}
	1	p_{10}	p_{11}	p_{1+}
		p_{+0}	p_{+1}	1

$\longrightarrow$ Is the rate of infection of virus A the same as that of virus B?

In other words: Is $p_{1+} = p_{+1}$? Equivalently, is $p_{10} = p_{01}$?

McNemar's test

H_0: $p_{01} = p_{10}$

Under H_0, e.g. if $p_{01} = p_{10}$, the expected counts for cells 01 and 10 are both equal to $(n_{01} + n_{10})/2$.

The χ^2 test statistic reduces to $X^2 = \dfrac{(n_{01} - n_{10})^2}{n_{01} + n_{10}}$

For large sample sizes, this statistic has null distribution that is approximately a χ^2(df = 1).

For the example: $X^2 = (20 - 9)^2 / 29 = 4.17 \quad \longrightarrow \quad P = 4.1\%$.

An exact test

Condition on $n_{01} + n_{10}$.

Under H_0, $n_{01} \mid n_{01} + n_{10} \sim \text{Binomial}(n_{01} + n_{10}, 1/2)$.

In R, use the function `binom.test`.

$\longrightarrow$ For the example, P = 6.1%.

Paired data

Paired data

	I-B	NI-B	
I-A	9	9	18
NI-A	20	62	82
	29	71	100

$\rightarrow$ P = 6.1%

Unpaired data

	I	NI	
A	18	82	100
B	29	71	100
	47	153	200

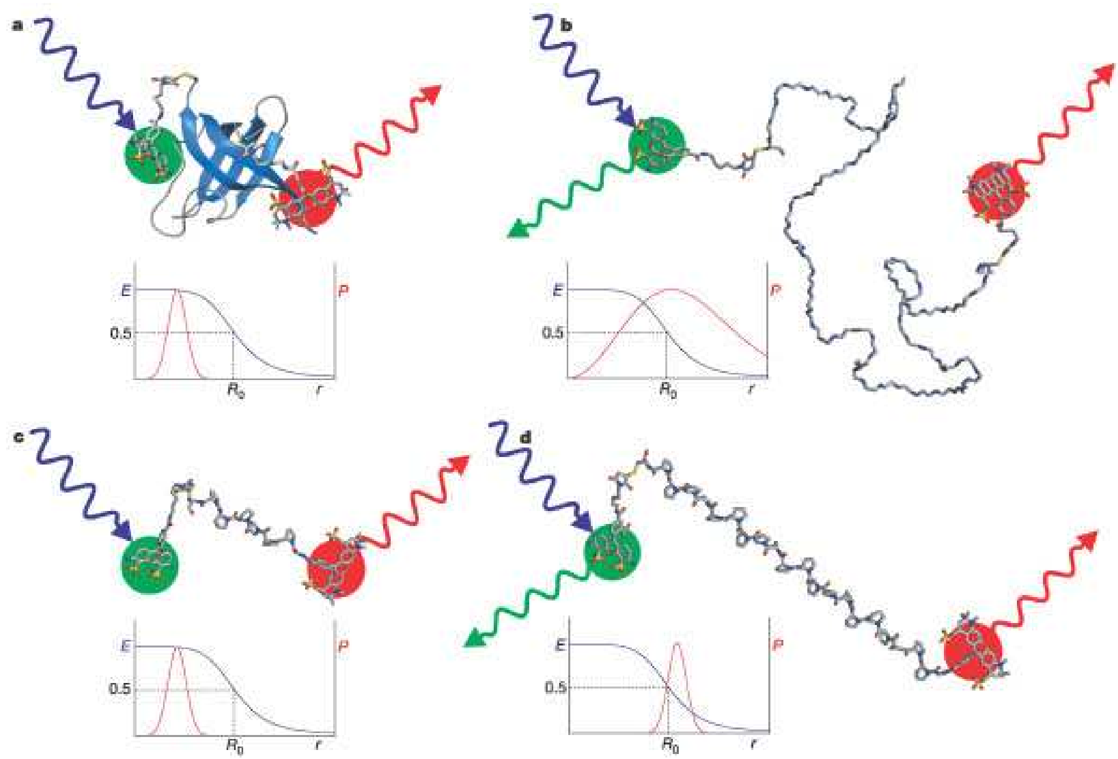
$\rightarrow$ P = 9.5%

$\longrightarrow$ Taking appropriate account of the “pairing” is important!

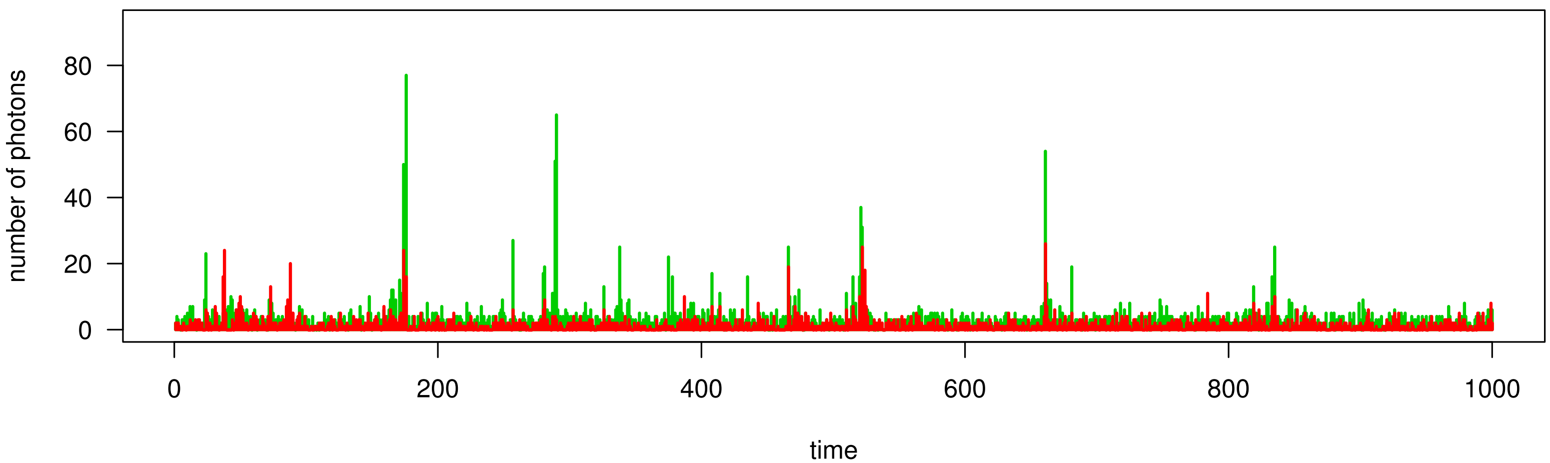
Deviations from Random Coil Behaviour

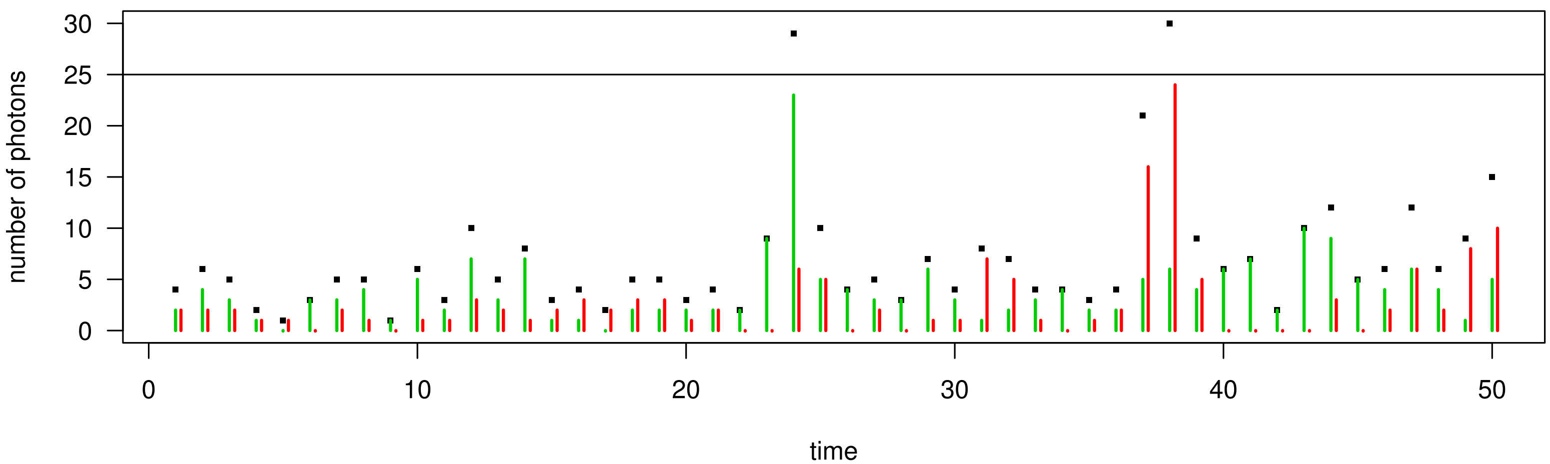
Are there site-specific deviations from random coil dimensions?

Förster Resonance Energy Transfer enables us to measure the distance between two dye molecules within a certain range. This can be used to study site-specific deviations from random coil dimensions in highly denatured peptides.

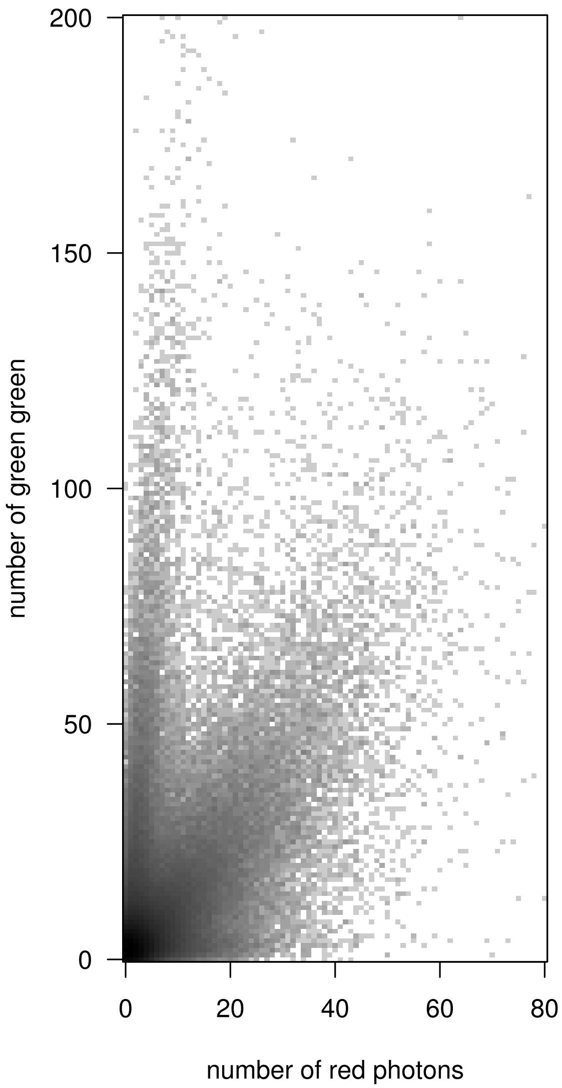


Deviations from Random Coil Behaviour





Deviations from Random Coil Behaviour

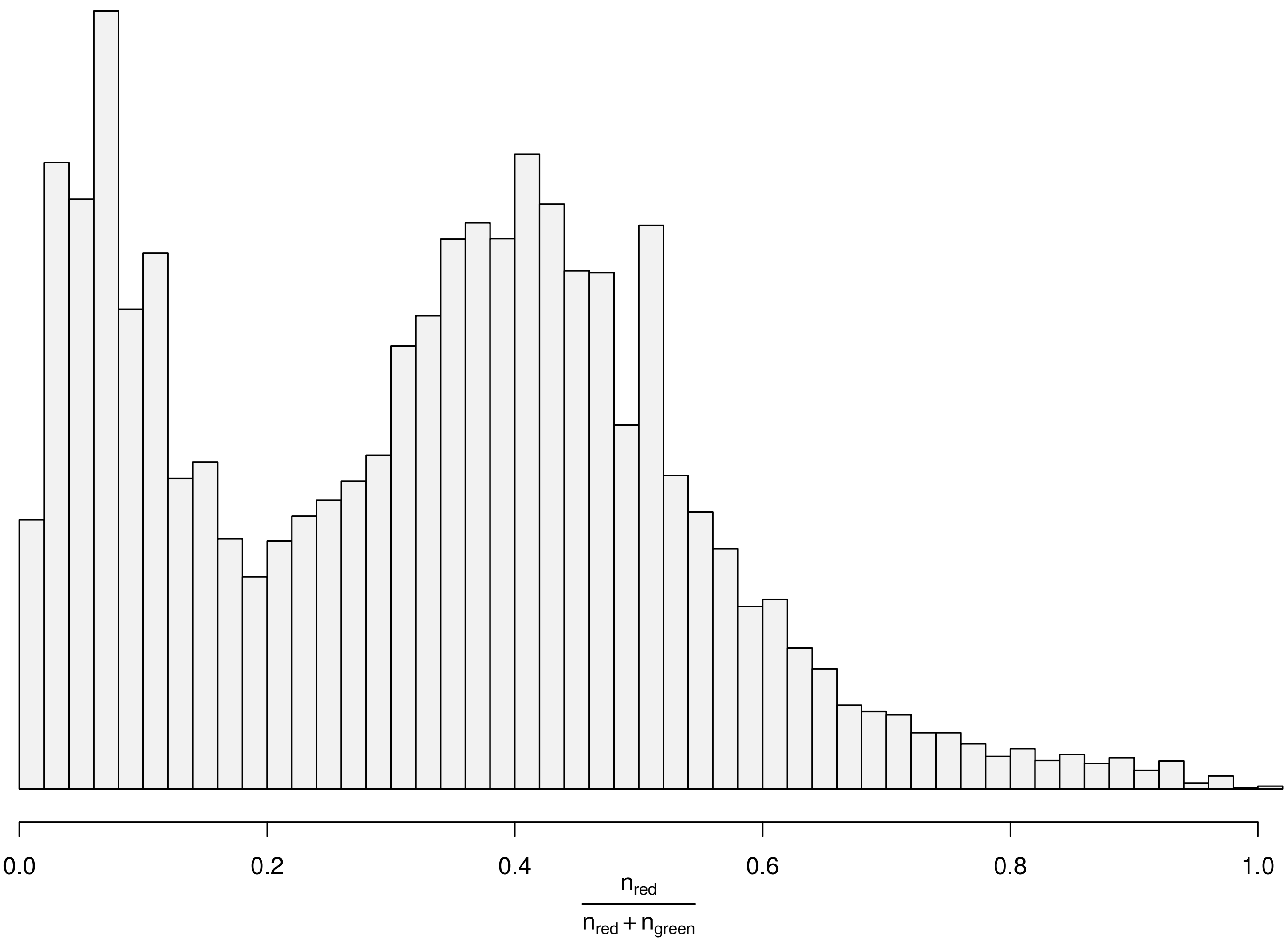


We have two underlying distributions for the green and red photons:

- One stemming from a peptide only having a donor dye.
- One stemming from a peptide being properly tagged with a donor and an acceptor dye.

Assume a photon has probability p_0 of being red in the former situation, and p_1 in the latter.

Deviations from Random Coil Behaviour



Deviations from Random Coil Behaviour

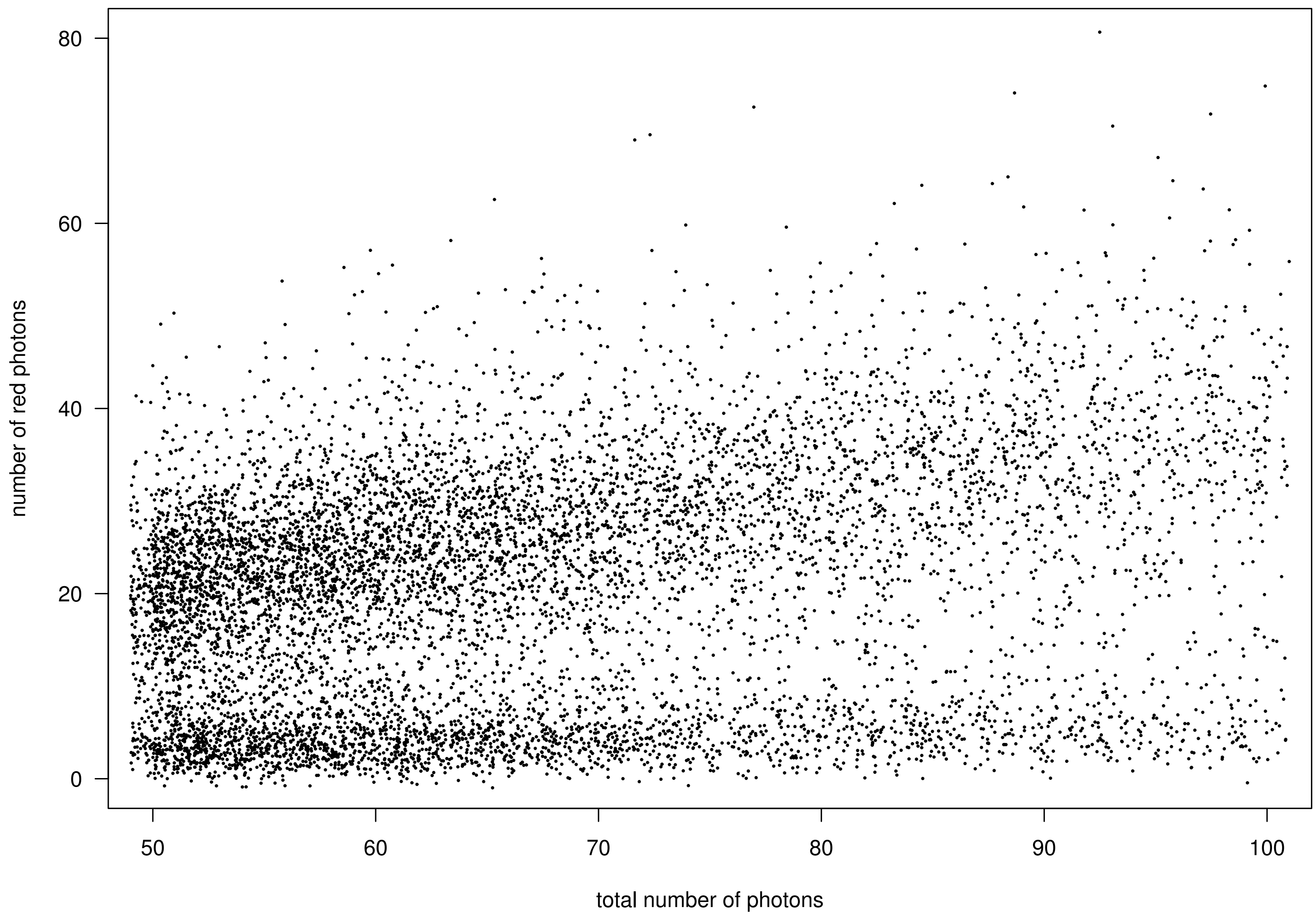
Assume we observe n_i photons at time point i. Then the number of red photons is simply Bernoulli(n_i, p_i), where p_i is either p_0 or p_1 . Assume that the probability of observing photons from a peptide without an acceptor dye at any time is p, independent of the total number of photons observed. Let X be the number of red photons. Then

$$\begin{aligned} P(X = x_i | n_i) &= P(X = x_i | n_i, p_0) \times p + P(X = x_i | n_i, p_1) \times (1 - p) \\ &= \binom{n_i}{x_i} p_0^{x_i} (1 - p_0)^{n_i - x_i} \times p + \binom{n_i}{x_i} p_1^{x_i} (1 - p_1)^{n_i - x_i} \times (1 - p), \end{aligned}$$

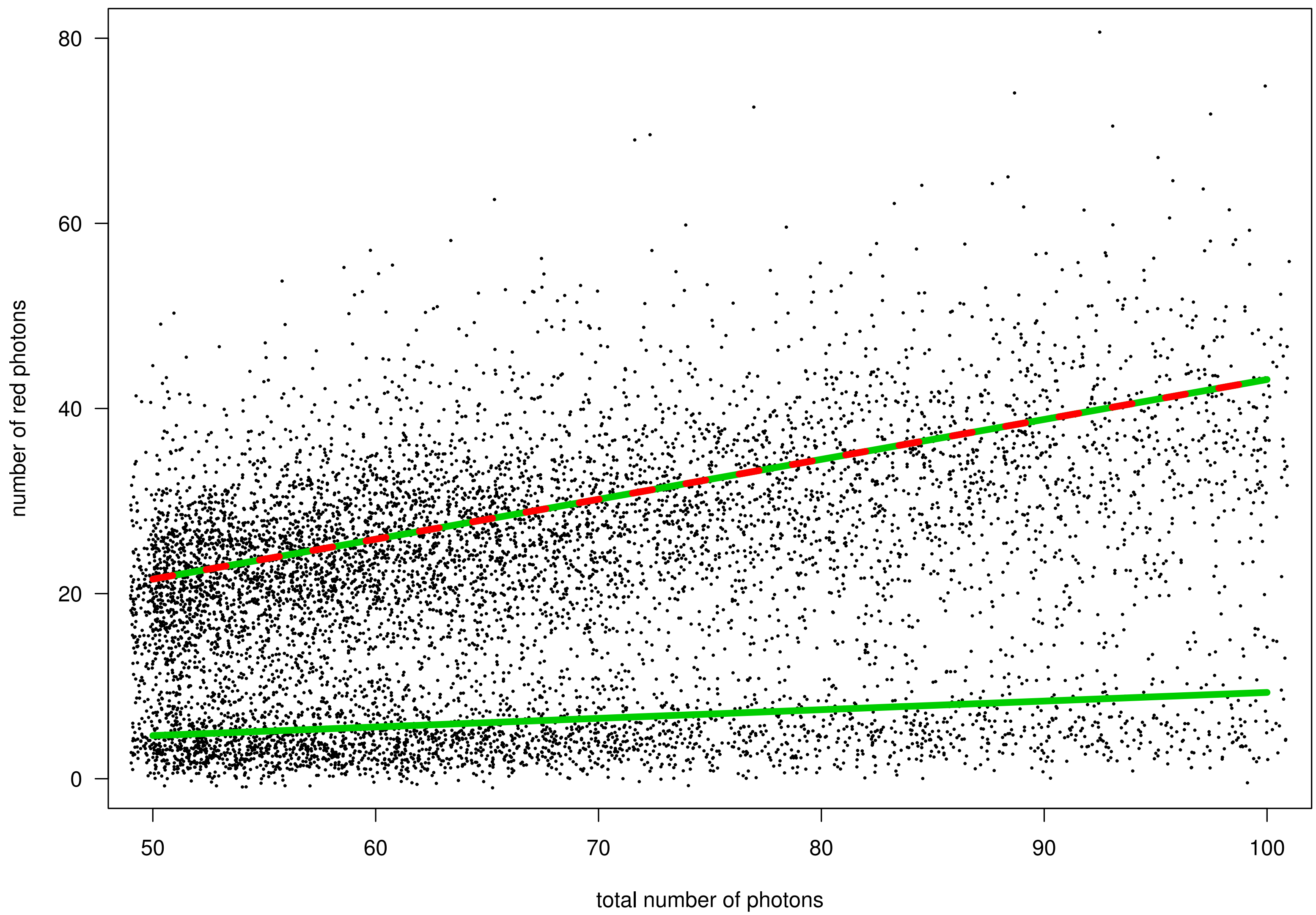
and hence

$$L(p, p_0, p_1) = \prod_{i=1}^{N} \left[\binom{n_i}{x_i} p_0^{x_i} (1 - p_0)^{n_i - x_i} \times p + \binom{n_i}{x_i} p_1^{x_i} (1 - p_1)^{n_i - x_i} \times (1 - p) \right].$$

Deviations from Random Coil Behaviour



Deviations from Random Coil Behaviour



Deviations from Random Coil Behaviour