

# Final: STAT 570

Philip Pham

December 7, 2018

Consider the failure time data in Table 1.

1. We describe a simple model for these data. Let  $p$  ( $0 < p < 1$ ) denote the weekly failure probability, i.e., the probability of failure during any week, and  $T$  the random variable describing the week at which failure occurred. Then  $T$  may be modeled as a geometric random variable:

$$\mathbb{P}(T = t | p) = \begin{cases} p(1-p)^{t-1}, & t = 1, 2, \dots; \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Let  $Y_t$  represent the number of components that fail in week  $t$ ,  $t = 1, 2, \dots, N$ , and  $Y_{N+1}$  the number of components that have not failed by week  $N$ .

- (a) Show that the likelihood function is

$$L(p) = \left[(1-p)^N\right]^{Y_{N+1}} \prod_{t=1}^N \left[p(1-p)^{t-1}\right]^{Y_t}. \quad (2)$$

**Solution:** An individual component's failure week has distribution Geometric( $p$ ).

The probability that a single component fails in week  $t$  is the probability that it survived  $t-1$  weeks and failed on week  $t$ , which is  $p(1-p)^{t-1}$ . There are  $Y_t$  such components, which gives us the factors for  $t = 1, 2, \dots, N$ .

The probability that a component fails at a later date is

$$(1-p)^N \sum_{k=1}^{\infty} p(1-p)^{k-1} = (1-p)^N \frac{p}{1-(1-p)} = (1-p)^N,$$

which gives us the remaining factor. There are  $Y_{N+1}$  remaining components, so

$$L(p) = \left\{ \prod_{t=1}^N \left[p(1-p)^{t-1}\right]^{Y_t} \right\} \times \left[(1-p)^N\right]^{Y_{N+1}}.$$

- (b) Find an expression for the MLE  $\hat{p}$ .

**Solution:** The score function is

$$\begin{aligned} S(p) &= \frac{\partial}{\partial p} \log L(p) \\ &= \frac{\partial}{\partial p} \left[ NY_{N+1} \log(1-p) + \sum_{t=1}^N Y_t (\log p + (t-1) \log(1-p)) \right] \\ &= -\frac{NY_{N+1}}{1-p} + \sum_{t=1}^N Y_t \left( \frac{1}{p} - \frac{t-1}{1-p} \right) = -\frac{NY_{N+1}}{1-p} + \sum_{t=1}^N Y_t \frac{1-pt}{p(1-p)}. \quad (3) \end{aligned}$$

Time (weeks), $i$	Failures, $y_i$	Temperature, $x_i$
1	210	24.0
2	108	26.0
3	58	24.0
4	40	26.0
5	17	25.0
6	10	22.0
7	7	23.0
8	6	20.0
9	5	21.0
10	4	18.0
11	2	17.0
12	3	20.0
> 12	15	

Table 1: Time until failure for  $n = 485$  components, along with average weekly temperature.

Solving for  $S(\hat{p}) = 0$ , we find the MLE:

$$\hat{p} \left( NY_{N+1} + \sum_{t=1}^N tY_t \right) = \sum_{t=1}^N Y_t \implies \boxed{\hat{p} = \frac{\sum_{t=1}^N Y_t}{NY_{N+1} + \sum_{t=1}^N tY_t}}. \quad (4)$$

- (c) Find the form of the observed information and hence the asymptotic variance of the MLE.

**Solution:** Using Equation 3, the expected observed information is

$$\begin{aligned}
I(p) &= \mathbb{E} \left[ -\frac{\partial}{\partial p} S(p) \mid p \right] \\
&= \frac{N \mathbb{E}[Y_{N+1} \mid p]}{(1-p)^2} + \sum_{t=1}^N \mathbb{E}[Y_t \mid p] \left( \frac{1}{p^2} + \frac{t-1}{(1-p)^2} \right) \\
&= n \frac{N(1-p)^N}{(1-p)^2} + np \sum_{t=1}^N (1-p)^{t-1} \left( \frac{1}{p^2} + \frac{t-1}{(1-p)^2} \right) \\
&= n \left[ \frac{(1-p)^N}{(1-p)^2} + \frac{1 - (1-p)^N}{p^2} + \frac{(1-p) - (1-p)^N}{p(1-p)^2} \right] \\
&= \boxed{n \frac{1 - (1-p)^N}{p^2(1-p)}}. \quad (5)
\end{aligned}$$

From Equation 5, the asymptotic variance of  $\hat{p}$  is

$$\text{var}(\hat{p}) \approx I(\hat{p})^{-1} = \frac{1}{n} \times \frac{p^2(1-p)}{1 - (1-p)^N}. \quad (6)$$