

# Final: STAT 570

Philip Pham

December 8, 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  $p(1-p)^{t-1}$  from Equation 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

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

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 maximum likelihood estimate (MLE).

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

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

where  $n = Y_{N+1} + \sum_{t=1}^N Y_t$ .

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

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

by asymptotic normality of the MLE.

- (d) For the data in Table 1, calculate the MLE,  $\hat{p}$ , the variance of  $\hat{p}$ , and an asymptotic 95% confidence interval for  $p$ .

**Solution:** The MLE can be calculated with Equation 4 to be  $\boxed{\hat{p} = 0.354717}$ .

The variance can be found with Equation 6 to be  $\boxed{\text{vâr}(\hat{p}) = 0.00016828}$ .

If  $\Phi$  is the cumulative distribution function for a standard normal, we can use asymptotic normality to find the 95% confidence interval as

$$\left[ \hat{p} + \Phi^{-1}(0.025) \sqrt{\text{vâr}(\hat{p})}, \hat{p} + \Phi^{-1}(0.975) \sqrt{\text{vâr}(\hat{p})} \right] = \boxed{[0.32929, 0.38014]}.$$

- (e) We now consider a Bayesian analysis. The conjugate prior for  $p$  is a beta distribution,  $\text{Beta}(a, b)$ . State the form of the posterior with this choice. Give the form of the posterior mean and write as a weighted combination of the MLE and the prior mean.

**Solution:** By Bayes' rule, we know the posterior density is proportional to the likelihood times the prior. From Equation 2, we'll have

$$\begin{aligned} L(p) \times [p^{a-1} (1-p)^{b-1}] &= p^{a-1} (1-p)^{b+NY_{N+1}-1} \prod_{t=1}^N [p(1-p)^{t-1}]^{Y_t} \\ &= p^{a+\sum_{t=1}^N Y_t-1} (1-p)^{b+\sum_{t=1}^N (t-1)Y_t+NY_{N+1}-1}, \end{aligned}$$

whose form we recognize as the integrand of beta function, so the posterior also has beta distribution, that is,

$$\begin{aligned} p \mid Y_1, Y_2, \dots, Y_{N+1} &\sim \text{Beta} \left( a + \sum_{t=1}^N Y_t, b + \sum_{t=1}^N (t-1)Y_t + NY_{N+1} \right) \\ &= \frac{\Gamma(a') \Gamma(b')}{\Gamma(a'+b')} p^{a'-1} (1-p)^{b'-1}, \end{aligned} \quad (7)$$

where  $a' = a + \sum_{t=1}^N Y_t$  and  $b' = b + \sum_{t=1}^N (t-1)Y_t + NY_{N+1}$ .

The posterior mean takes the form

$$\begin{aligned} \mathbb{E}[p \mid Y_1, Y_2, \dots, Y_{N+1}] &= \frac{a'}{a'+b'} \\ &= \frac{a + \sum_{t=1}^N Y_t}{a + b + \sum_{t=1}^N tY_t + NY_{N+1}}. \end{aligned} \quad (8)$$

We have that the prior mean is  $p_{\text{prior}} = \frac{a}{a+b}$ . Equation 8 can be rewritten as

$$\boxed{\frac{(a+b)p_{\text{prior}} + \left(\sum_{t=1}^N tY_t + NY_{N+1}\right)\hat{p}}{a+b+\sum_{t=1}^N tY_t + NY_{N+1}}}, \quad (9)$$

so the posterior mean is a convex combination of the prior mean and MLE.

- (f) Suppose we wish to fix the parameters of the prior,  $a$  and  $b$ , so that the mean is  $\mu$  and the prior standard deviation is  $\sigma$ . Obtain expressions for  $a$  and  $b$  in terms of  $\mu$  and  $\sigma^2$ .

**Solution:** It is well known that the mean and variance of the  $\text{Beta}(a, b)$  distribution are  $\frac{a}{a+b}$  and  $\frac{ab}{(a+b)^2(a+b+1)}$ , respectively. Solving equations

$$\begin{aligned} \frac{a}{a+b} &= \mu \\ \frac{ab}{(a+b)^2(a+b+1)} &= \sigma^2, \end{aligned}$$

we find that

$$a = \mu \left[ \frac{\mu(1-\mu)}{\sigma^2} - 1 \right] \quad (10)$$

$$b = (1-\mu) \left[ \frac{\mu(1-\mu)}{\sigma^2} - 1 \right]. \quad (11)$$

- (g) For the data in Table 1, assume we wish to have a beta prior with  $\mu = 0.2$  and  $\sigma = 0.08$ . State the posterior for the prior corresponding to this choice and evaluate the posterior mean. Simulate samples from the posterior distribution. Provide a histogram representation of the posterior distribution and calculate the 5%, 50% and 95% points of the posterior distribution.

**Solution:**

2. (a) A more complex likelihood for these data would assume that the  $i$ -th component had their own probability  $p_i$ , with the  $p_i$ 's arising from a distribution  $\pi(p)$ . Show that

$$\mathbb{P}(T = t) = \mathbb{E}[(1 - p)^{t-1}] - \mathbb{E}[(1 - p)^t], \quad (12)$$

and

$$\mathbb{P}(T > t) = \mathbb{E}[(1 - p)^t]. \quad (13)$$

**Solution:** First let us find the survival function in 13.

$$\begin{aligned} \mathbb{P}(T > t) &= \int_0^1 \mathbb{P}(T > t \mid p) \pi(p) \, dp = \int_0^1 \left[ \sum_{s=t+1}^{\infty} p(1-p)^{s-1} \right] \pi(p) \, dp \\ &= \int_0^1 \left[ p \sum_{s=0}^{\infty} (1-p)^s \right] (1-p)^t \pi(p) \, dp \\ &= \int_0^1 \left[ p \times \frac{1}{1 - (1-p)} \right] (1-p)^t \pi(p) \, dp = \int_0^1 (1-p)^t \pi(p) \, dp \\ &= \mathbb{E}[(1-p)^t], \end{aligned}$$

which proves Equation 13.

The probability mass function in Equation 12 follows:

$$\mathbb{P}(T = t) = \mathbb{P}(T > t - 1) - \mathbb{P}(T > t) = \mathbb{E}[(1-p)^{t-1}] - \mathbb{E}[(1-p)^t].$$

- (b) Obtain expressions for  $\mathbb{P}(T = t \mid \alpha, \beta)$  and  $\mathbb{P}(T > t \mid \alpha, \beta)$  with  $\pi(\cdot)$  taken as the beta distribution,  $\text{Beta}(\alpha, \beta)$ .

**Solution:** These follow from Equations 12 and 13.

$$\begin{aligned} \mathbb{P}(T > t) &= \mathbb{E}[(1-p)^t] = \sum_{s=t}^{\infty} \mathbb{E}[p(1-p)^s] \\ &= \sum_{s=t}^{\infty} \int_0^p p(1-p)^s \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1} (1-p)^{\beta-1} \, dp \\ &= \sum_{s=t}^{\infty} \int_0^p \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha+1-1} (1-p)^{\beta+s-1} \, dp \\ &= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \sum_{s=t}^{\infty} \frac{\Gamma(\alpha + 1)\Gamma(\beta + s)}{\Gamma(\alpha + \beta + s + 1)} = \alpha \frac{\Gamma(\alpha + \beta)}{\Gamma(\beta)} \sum_{s=t}^{\infty} \frac{\Gamma(\beta + s)}{\Gamma(\alpha + \beta + s + 1)} \\ &= 1 - \alpha \frac{\Gamma(\alpha + \beta)}{\Gamma(\beta)} \sum_{s=0}^{t-1} \frac{\Gamma(\beta + s)}{\Gamma(\alpha + \beta + s + 1)} \\ &= 1 - \frac{1}{B(\alpha, \beta)} \sum_{s=0}^{t-1} B(\alpha + 1, \beta + s) = \frac{B(\alpha, \beta + t)}{B(\alpha, \beta)} = \frac{\Gamma(\alpha + \beta)\Gamma(\beta + t)}{\Gamma(\beta)\Gamma(\alpha + \beta + t)} \end{aligned} \quad (14)$$

where  $B$  is the beta function, and we know  $\mathbb{P}(T > 0) = 1$ .

Plugging Equation 14 into Equation 12, one obtains

$$\mathbb{P}(T = t) = \frac{B(\alpha + 1, \beta + t - 1)}{B(\alpha, \beta)} = \alpha \frac{\Gamma(\alpha + \beta) \Gamma(\beta + t - 1)}{\Gamma(\beta) \Gamma(\alpha + \beta + t)} \quad (15)$$

for  $t \in \mathbb{N}$ .

- (c) Using the previous part, write down the likelihood function  $L(\alpha, \beta)$  corresponding to data  $\{Y_t\}_{t=1}^{N+1}$ .

**Solution:** Our model for  $T$  is different, so we can substitute Equations 15 and 14 into Equation 2: we'll have  $\mathbb{P}(T = t)$  in place of  $p(1 - p)^{t-1}$  and  $\mathbb{P}(T > N)$  in place of  $(1 - p)^N$ .

$$\begin{aligned} L(\alpha, \beta) &= [\mathbb{P}(T > N)]^{Y_{N+1}} \prod_{t=1}^N [\mathbb{P}(T = t)]^{Y_t} \\ &= \left[ \frac{B(\alpha, \beta + N)}{B(\alpha, \beta)} \right]^{Y_{N+1}} \prod_{t=1}^N \left[ \frac{B(\alpha + 1, \beta + t - 1)}{B(\alpha, \beta)} \right]^{Y_t}. \end{aligned} \quad (16)$$

- (d) Find the MLEs  $\hat{\alpha}$  and  $\hat{\beta}$  for the data of Table 1.

**Solution:** From Equation 16, we can consider the log-likelihood function:

$$\begin{aligned} l(\alpha, \beta) &= \log L(\alpha, \beta) \\ &= -n \log B(\alpha, \beta) + Y_{N+1} \log B(\alpha, \beta + N) + \sum_{t=1}^N Y_t \log B(\alpha + 1, \beta + t - 1). \end{aligned} \quad (17)$$

The score function is

$$\begin{aligned} S(\alpha, \beta) &= \nabla l(\alpha, \beta) = \begin{pmatrix} S_\alpha(\alpha, \beta) \\ S_\beta(\alpha, \beta) \end{pmatrix} \\ S_\alpha(\alpha, \beta) &= -n [\psi(\alpha) - \psi(\alpha + \beta)] + Y_{N+1} [\psi(\alpha) - \psi(\alpha + \beta + N)] \\ &\quad + \sum_{t=1}^N Y_t [\psi(\alpha + 1) - \psi(\alpha + \beta + t)], \\ S_\beta(\alpha, \beta) &= -n [\psi(\beta) - \psi(\alpha + \beta)] + Y_{N+1} [\psi(\beta + N) - \psi(\alpha + \beta + N)] \\ &\quad + \sum_{t=1}^N Y_t [\psi(\beta + t - 1) - \psi(\alpha + \beta + t)], \end{aligned} \quad (18)$$

where  $\psi(x) = \Gamma'(x) / \Gamma(x)$  is the digamma function.

Numerically solving Equation 18 for  $S(\hat{\alpha}, \hat{\beta}) = \mathbf{0}$ , I obtain  $\hat{\alpha} = 1.413336$

and  $\hat{\beta} = 1.38001102$  for the MLEs.

3. (a) Show that the likelihood in Equation 1 can be written as a product of binomial distributions, the form of which you should precisely give.

**Solution:**

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.