MS&E 121
Introduction to Stochastic Modeling
Prof. Peter W. Glynn

Motivating Problem

We are members of a design team that has been tasked to create a new component for a satellite that will go into a "near earth" orbit (with a mission duration of two weeks).

The component includes a new capacitor. The capacitor has been tested experimentally, resulting in lifetimes of: 3, 1, 5, 2, and 1 months.

Question: What is the probability that the capacitor will last more than two weeks? (1/2 month)

Analysis: All the experimental observations are greater then or equal to 1/2 month. However, the sample size ($n = 5$) is too small to be fully representative. We know intuitively that there is a positive probability of failure in the first half month. So, we need a model.

Model: Previous version of the capacitors (for which lots of data is available) showed that the exponential distribution fit the lifetime data well. We will assume an exponential distribution for the lifetime of our new capacitor.

In other words, let T = lifetime of capacitor. We are assuming that the probability density function of T takes the form

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t} & , t \geq 0 \\ 0 & , t < 0. \end{cases}$$

How should we fit the unknown parameter λ from the observed data? We will use the most important of all statistical estimation principles.

Principle of Maximum Likelihood

Estimate λ via the value $\widehat{\lambda}$ of the parameter that maximizes the likelihood of the observed sample (in this case, 3, 1, 5, 2, 1).

Under parameter λ, the likelihood of the first observation is $\lambda\exp(-3\lambda)$. The likelihoods for the other four observations are $\lambda\exp(-\lambda)$, $\lambda\exp(-5\lambda)$, $\lambda\exp(-2\lambda)$, and $\lambda\exp(-\lambda)$, respectively. We can (reasonably) expect the five experimental outcomes to be independent, so the likelihood for the entire sample is

$$\begin{aligned} \mathrm{L}(\lambda) &= \lambda\exp(-3\lambda)\lambda\exp(-\lambda)\lambda\exp(-5\lambda)\lambda\exp(-2\lambda)\lambda\exp(-\lambda) \\ &= \lambda^5\exp(-12\lambda) \end{aligned}$$

What value $\widehat{\lambda}$ maximizes the likelihood of the observed sample?

Maximizing the likelihood is the same as maximizing the log-likelihood.

$$\begin{aligned} \mathcal{L} &= \log L(\lambda) \\ &= 5 \log \lambda - 12\lambda \end{aligned}$$

At a maximizer $\widehat{\lambda}$,

$$\frac{5}{\widehat{\lambda}} = 12.$$

So, the maximum likelihood estimator for the "true parameter" λ^* is

$$\widehat{\lambda} = \frac{5}{12}.$$

Doing the calculation: Now that we have an estimator for the parameter (as computed from the data), we can compute $P(T > 1/2)$. For an exponential(λ) random variable (rv),

$$P(T > 1/2) = \int_{1/2}^{\infty} \lambda e^{-\lambda t} = \exp(\lambda/2)$$

Plugging in our estimator $\widehat{\lambda}$, our answer for the probability that the component does not fail during the mission is exp(-5/24).

Problem Modification 1: Suppose that we have six observations for the lifetime data : 3, 1, 5, 2, 1, and we are told the sixth component is still functioning seven months after having started the test. How does this affect our analysis?

The estimator $\widehat{\lambda}$ changes. The likelihood that an exponential rv having parameter λ survives more than 7 month os exp(-7λ). So, the likelihood of the sample of six observations is

$$L(\lambda) = \lambda^5 \exp(-19\lambda).$$

The maximum likelihood estimator is

$$\widehat{\lambda} = \frac{5}{19},$$

so our answer for the probability that the component does not fail during the mission is exp(-5/38).

Problem Modification 2: The satellite component now consists of two capacitor in series :

If T_1, T_2 are respectively the lifetimes of the two capacitors, the system lifetime is $T = \min(T_1, T_2)$. Assume the data collected on the capacitor lifetime is as before : 3, 1, 5, 2, 1.

Doing the calculation: Note that

$$\begin{aligned} P(T > t) &= P(\min(T_1, T_2) > t) \\ &= P(T_1 > t, T_2 > t). \end{aligned}$$

Assume the two capacitors fail independently. Then,

$$\begin{aligned} P(T_1 > t, T_2 > t) &= P(T_1 > t)P(T_2 > t) \\ &= \exp(-\lambda t)\exp(-\lambda t) \\ &= \exp(-2\lambda t). \end{aligned}$$

But our estimator for the true λ^* is $\widehat{\lambda} = 5/12$. So, our estimate for the system surviving 1/2 month is $\exp(-5/12)$.

Problem Modification 3: Is it really the case that the two capacitors fail independently? There is a passibility of a "common node" failure, caused by a voltage surge through the system. If such a voltage surge occurs, then both capacitor fail simultaneously.

To model this, let Z be the (random) time during the mission at which the voltage surge occurs. Our model for the "common node" failure is :

$$P(T > t | Z = z) = \begin{cases} \exp(-2\lambda t) & , t < z \\ 0 & , t \geq 0. \end{cases}$$

Doing the calculation: We compute the distribution of system lifetime by conditioning :

$$\begin{aligned} P(T > t) &= \int_0^\infty P(T > t | Z = z) f_Z(z) dz \\ &= \int_t^\infty P(T > t | Z = z) f_Z(z) dz \\ &= \exp(-2\lambda t) P(Z > t). \end{aligned}$$

Again, our estimator $\widehat{\lambda} = 1/2$, so we estimate the probability of survival for more than 1/2 month as $\exp(-5/12)P(Z > 1/2)$. ($P(Z > 1/2)$ would need to come from members of the development team dealing with electrical surge issues.)

Problem Modification 4: In this design, the two capacitors are in parallel:

here, $T = \max(T_1, T_2)$.

Doing the calculation: Assuming the two capacitors fail independently,

$$\begin{aligned} P(T \leq t) &= P(\max(T_1, T_2) \leq t) \\ &= P(T_1 \leq t, T_2 \leq t) \\ &= (1 - \exp(-\lambda t))(1 - \exp(-\lambda t)). \end{aligned}$$

So,

$$P(T > t) = 1 - (1 - \exp(-\lambda t))(1 - \exp(-\lambda t)).$$

Our estimator $\widehat{\lambda} = 5/12$ and $t = 1/2$, so our estimated probability is

$$1 - (1 - \exp(-5/24))(1 - \exp(-5/24))$$

.

Problem Modification 5: We now consider a design involving two capacitors, in which one is kept on "cold stand-by". The circuit involving the "cold stand-by" unit is inly activated when the other circuit fails. in this case, the total system lifetime is $T = T_1 + T_2$.

Doing the calculation: If T_1, T_2 are dependent rv's, we must compute $P(T_1 + T_2 > t)$ by integrating against the joint probability density function :

$$P(T_1 + T_2 > t) = \int_0^t \int_{t-s}^\infty f_{T_1,T_2}(s, u) du ds + \int_t^\infty \int_0^\infty f_{T_1,T_2}(s, u) du ds.$$

Given our cold stand-by setting, it is reasonable to assume that T_1, T_2 are independent, so that

$$
\begin{aligned}
P(T_1+T_2>t) &= \int_0^t \int_{t-s}^\infty \lambda^2 e^{-\lambda(s+u)} du ds + \int_t^\infty \int_0^\infty \lambda^2 e^{\lambda(s+u)} du ds \\
&= e^{-\lambda t} + \lambda t e^{-\lambda t}.
\end{aligned}
$$

Our estimator $\widehat{\lambda} = 5/12$, so our estimate of survival for more than 1/2 month is $\exp(-5/24)(29/24)$.

Problem Modification 6: We are going to manufacture large number of Design 1. The design has now been engineered so that $P(T > 1/2) = 0.95$. We will be testing all outgoing components. If the component is "good" (i.e. it will have a lifetime greater then 1/2 month), it passes the test. If it is "bad" (has a lifetime less than 1/2 month), it fails the test with probability 0.9.

Question: What proportion of outgoing components is "bad"?

$$
\begin{aligned}
& P(\text{ a component sent out is bad }) \\
=\; & P(\text{ component sent out } \mid \text{ bad })P(\text{ bad }) \\
=\; & \tfrac{1}{10} \cdot \tfrac{1}{20} = \tfrac{1}{200}.
\end{aligned}
$$

So, if we send out 1000 components to a customer, we expect roughly 5 of them to be bad. The customer does more extensive testing at their facility; so can determine precisely which components are good/bad. To ensure that the customer gets 1000 good components, we decide to ship 1005 components to the customer.

Question: What is the probability that the customer receives at least 1000 good components?

Let N = number of good components in shipment of 1005 components. If each component is independently good/bad (a reasonable assumption?),

$$N \stackrel{\mathcal{D}}{=} \text{binomial}\Big(1005, \frac{199}{200}\Big).$$

(Here, $\stackrel{\mathcal{D}}{=}$ means "has the distribution of"). So,

$$P(N \geq 1000) = \sum_{j=1000}^{1005} \binom{1005}{j} \Big(\frac{199}{200}\Big)^j \Big(\frac{1}{200}\Big)^{1005-j}$$

A Useful Approximation: If $\widetilde{N} \stackrel{\mathcal{D}}{=} \text{binomial}(n,p)$ where n is big and p is small, then

$$\widetilde{N} \stackrel{\mathcal{D}}{\approx} \text{Poisson}(np)$$

(where $\stackrel{\mathcal{D}}{\approx}$ means "has approximately the distribution of"). How can we use this in the above problem?

Let $\widetilde{N}$ = number of bad components in shipment of size 1005. Then

$$\widetilde{N} = \text{binomial}\Big(1005, \frac{1}{200}\Big).$$

We can therefore apply our approximation to $\widetilde{N}$. So, $P(N \geq 1000) = P(\widetilde{N} \leq 5)$ and

$$P(\widetilde{N} \leq 5) = \sum_{j=0}^{5} e^{-\frac{1005}{200}} \Big(\frac{1005}{200}\Big)^j \frac{1}{j!}.$$