UC Berkeley
Department of Electrical Engineering and Computer Sciences

ELECTRICAL ENGINEERING 126: PROBABILITY AND RANDOM PROCESSES

Midterm 2 Review
Fall 2017

---

1. **Two-Population Sampling**

We are conducting a public opinion poll to determine the fraction p of people who will vote for Mr. Whatshisname as the next president. We ask N_1 college-educated and N_2 non-college-educated people, where N_1 and N_2 are positive integers. We assume that the votes in each of the two groups are i.i.d. Bernoulli(p_1) and Bernoulli(p_2), respectively in favor of Whatshisname. In the general population, the percentage of college-educated people is known to be q.

(a) What is a 95% confidence interval for p, using an upper bound for the variance?

(b) How do we choose N_1 and N_2 subject to $N_1 + N_2 = N$ to minimize the width of that interval? (You may ignore the constraint that N_1 and N_2 must be integers.)

Solution:

(a) If we let $\hat{p}_1$ and $\hat{p}_2$ denote the fraction of people who vote for Mr. Whatshisname in the two groups respectively, then an unbiased estimator for p is $\hat{p} := q\hat{p}_1 + (1-q)\hat{p}_2$. The variance of $\hat{p}$ is

$$\operatorname{var}\hat{p} = \frac{q^2 p_1(1-p_1)}{N_1} + \frac{(1-q)^2 p_2(1-p_2)}{N_2} \le \frac{1}{4}\Big(\frac{q^2}{N_1} + \frac{(1-q)^2}{N_2}\Big).$$

So, an approximate 95% confidence interval for p, using the CLT, is $\hat{p} \pm \sqrt{q^2/N_1 + (1-q)^2/N_2}$.

(b) Minimizing the width of the interval is equivalent to minimizing the variance. We can explicitly enforce the constraint by writing $N_2 = N - N_1$, and then we have:

$$\frac{\mathrm{d}}{\mathrm{d}N_1}\Big(\frac{q^2}{N_1} + \frac{(1-q)^2}{N-N_1}\Big) = -\frac{q^2}{N_1^2} + \frac{(1-q)^2}{(N-N_1)^2}.$$

The second derivative is positive so the function is convex, and so the first-order condition tells us the minimizer. Setting the derivative to 0, we find that $q/N_1 = (1-q)/(N-N_1)$. Therefore, the minimizer is $N_1 = qN$, $N_2 = (1-q)N$.

2. **Minimum and Maximum of Exponentials**

Let $\lambda_1, \lambda_2 > 0$, and $X_1 \sim \text{Exponential}(\lambda_1)$, $X_2 \sim \text{Exponential}(\lambda_2)$ are independent. Also, define $U := \min(X_1, X_2)$ and $V := \max(X_1, X_2)$. Show that U and $V - U$ are independent.

Solution:

For $u, w > 0$,

$$\begin{aligned}
\mathbb{P}(U \le u, V - U \le w, X_1 < X_2) &= \mathbb{P}(X_1 \le u, X_1 < X_2 \le X_1 + w) \\
&= \int_0^u \int_{x_1}^{x_1+w} \lambda_2 \exp(-\lambda_2 x_2)\,\mathrm{d}x_2\, \lambda_1 \exp(-\lambda_1 x_1)\,\mathrm{d}x_1 \\
&= \int_0^u \bigl\{\exp(-\lambda_2 x_1) - \exp\bigl(-\lambda_2(x_1 + w)\bigr)\bigr\}\lambda_1 \exp(-\lambda_1 x_1)\,\mathrm{d}x_1 \\
&= \bigl(1 - \exp(-\lambda_2 w)\bigr) \int_0^u \lambda_1 \exp\bigl(-(\lambda_1 + \lambda_2)x_1\bigr)\,\mathrm{d}x_1 \\
&= \frac{\lambda_1}{\lambda_1 + \lambda_2}\bigl(1 - \exp\{-(\lambda_1 + \lambda_2)u\}\bigr)\bigl(1 - \exp(-\lambda_2 w)\bigr).
\end{aligned}$$

By symmetry, interchanging the roles of 1 and 2 yields

$$\begin{aligned}
&\mathbb{P}(U \le u, V - U \le w, X_2 < X_1) \\
&\qquad = \frac{\lambda_2}{\lambda_1 + \lambda_2}\bigl(1 - \exp\{-(\lambda_1 + \lambda_2)u\}\bigr)\bigl(1 - \exp(-\lambda_1 w)\bigr).
\end{aligned}$$

Adding these two expressions yields

$$\begin{aligned}
\mathbb{P}(U \le u, V - U \le w) &= \bigl(1 - \exp\{-(\lambda_1 + \lambda_2)u\}\bigr)p_w, \qquad \text{where} \\
p_w &:= \frac{\lambda_1}{\lambda_1 + \lambda_2}\bigl(1 - \exp(-\lambda_2 w)\bigr) + \frac{\lambda_2}{\lambda_1 + \lambda_2}\bigl(1 - \exp(-\lambda_1 w)\bigr).
\end{aligned}$$

The joint CDF splits into a product of factors $\mathbb{P}(U \le u)\mathbb{P}(V - U \le w)$ which proves independence. To interpret the second term, observe that $\lambda_1/(\lambda_1 + \lambda_2)$ is the probability of the event $\{X_1 < X_2\}$; and conditioned on this event, $V - U \sim \text{Exponential}(\lambda_2)$ by the memoryless property.

3. **Integrated Shot Noise**

A noise impulse occurs at time $t = 0$, and later impulses occur at Poisson process times with mean rate $\lambda > 0$. Each impulse instantaneously charges a capacitor to 1 volt, and the voltage then decreases exponentially as e^{-t} until the next impulse occurs. Let V_t denote the voltage at time t. Let

$$Z_n = \int_0^{T_n} V_t\,\mathrm{d}t$$

be the integrated voltage up to the time of the n^{th} impulse occurring after $t = 0$, for each positive integer n.

Find $\mathbb{E}[Z_n]$ and $\operatorname{var} Z_n$.

Solution:

At T_n, there have been n impulses, and the interarrival times are i.i.d. and exponentially distributed. Thus,

$$\begin{aligned}
\mathbb{E}[Z_n] &= \mathbb{E}\Big[\int_0^{T_n} V_t \,\mathrm{d}t\Big] = \sum_{i=1}^{n} \mathbb{E}\Big[\int_{T_{i-1}}^{T_i} V_t \,\mathrm{d}t\Big] \\
&= n\int_0^\infty \int_0^s \exp(-t)\lambda\exp(-\lambda s)\,\mathrm{d}t\,\mathrm{d}s \\
&= n\int_0^\infty \big(1-\exp(-s)\big)\lambda\exp(-\lambda s)\,\mathrm{d}s = n\Big(1-\frac{\lambda}{1+\lambda}\Big) = \frac{n}{1+\lambda}
\end{aligned}$$

and

$$\begin{aligned}
\operatorname{var} Z_n &= \sum_{i=1}^{n} \operatorname{var}\int_{T_{i-1}}^{T_i} V_t\,\mathrm{d}t \\
&= n\Big\{\int_0^\infty \Big(\int_0^s \exp(-t)\,\mathrm{d}t\Big)^2 \lambda\exp(-\lambda s)\,\mathrm{d}s - \frac{1}{(1+\lambda)^2}\Big\} \\
&= n\Big(\int_0^\infty \big(1-\exp(-s)\big)^2\lambda\exp(-\lambda s)\,\mathrm{d}s - \frac{1}{(1+\lambda)^2}\Big) \\
&= n\Big(1-\frac{2\lambda}{1+\lambda}+\frac{\lambda}{2+\lambda}-\frac{1}{(1+\lambda)^2}\Big) = \frac{n\lambda}{(1+\lambda)^2(2+\lambda)}.
\end{aligned}$$

4. **Doubly Stochastic Matrix**

A matrix is called **doubly stochastic** if all of its entries are non-negative and each row and each column sums to 1. Find the stationary distribution for a doubly stochastic irreducible matrix.

Solution:

The stationary distribution is uniform over the state space $\mathcal{X}$. Indeed, if we define $\pi(x) := |\mathcal{X}|^{-1}$ for each $x \in \mathcal{X}$, then

$$\sum_{y\in\mathcal{X}} \pi(y)P(y,x) = |\mathcal{X}|^{-1}\sum_{y\in\mathcal{X}} P(y,x) = |\mathcal{X}|^{-1} = \pi(x).$$

5. **Flea on a Triangle**

A flea hops about at random on the vertices of a triangle, with all jumps equally likely. Find the probability that after n ($n \in \mathbb{N}$) hops the flea is back where it started.

Solution:

Let x be the flea's starting state. When $n = 0$, $p^{(0)}(x,x) = 1$, and when $n = 1$, $p^{(1)}(x,x) = 0$. Otherwise, for $n \in \mathbb{Z}_{\geq 2}$, in order for the flea to return to its original state after n hops, it must have *not* returned to its original state after $n-1$ hops, and then return to its original state. So, we have the recurrence $p^{(n)}(x,x) = (1-p^{(n-1)}(x,x))/2$. Iterating this recurrence, we find that

$$p^{(n)}(x,x) = \frac{1-p^{(n-1)}(x,x)}{2} = \frac{1}{2}\sum_{k=0}^{n-2}\Big(-\frac{1}{2}\Big)^k + \frac{(-1)^{n-1}}{2^n} + \frac{(-1)^n}{2^n}p^{(0)}(x,x)$$

$$= \frac{1}{2}\sum_{k=0}^{n-2}\Big(-\frac{1}{2}\Big)^k = \frac{(1/2)(1-(-1/2)^{n-1})}{3/2} = \frac{1-(-1/2)^{n-1}}{3}.$$