Lecture 14

§1 Survival function and hazard rate function

令 X 为一个正连续型随机变量，我们将其视作某物的 lifetime. X 有分布函数 F 与概率密度函数 f

1. **Definition:** survival function（生存/残存函数）& hazard rate function（危险率函数）

关于 X 的 survival function 被定义为

$$S(t) = \Pr(X > t) = 1 - F(t) \quad (\text{即 tail probability})$$

它表示了物体能至少存活 t 时间的概率.

而 hazard rate function（mortality function，failure rate function）被定义为

$$\lambda(t) = \frac{f(t)}{S(t)}$$

2. hazard rate function 的实际意义

The hazard rate function $\lambda(t)$ 表示了 t-unit-old 的物体会在 t 时刻突然死亡的 conditional probability intensity（注意 hazard rate 不是概率）

$$\Pr(t < X < t+dt \mid X > t) = \frac{F(t+dt) - F(t)}{S(t)}$$

$$\approx \frac{f(t)\,d(t)}{S(t)}$$

$$= \lambda(t)\,d(t)$$

3. **Property:** $\lambda(t)$ 与 F 唯一对应

The hazard rate function $\lambda(t)$ 唯一确定分布函数 $F(t)$

$$\lambda(t) = \frac{f(t)}{S(t)} = \frac{F'(t)}{S(t)} = -\frac{S'(t)}{S(t)} = -\frac{d}{dt}[\ln S(t)]$$

$$\Rightarrow \int_0^x \lambda(t)\,d(t) = -\ln S(t)\Big|_0^x = -\ln S(x) = -\ln[1 - F(x)]$$

$$\Rightarrow F(x) = 1 - e^{-\int_0^x \lambda(t)dt}$$

注：分布函数，生存函数，密度函数，危险率函数均可唯一确定分布

Functions in terms of	$F(t)$	$S(t)$	$f(t)$	$\lambda(t)$
$F(t) =$	$F(t)$	$1 - S(t)$	$\int_0^t f(s)\mathrm{d}s$	$1 - e^{-\int_0^t \lambda(s)\mathrm{d}s}$
$S(t) =$	$1 - F(t)$	$S(t)$	$\int_t^\infty f(s)\mathrm{d}s$	$e^{-\int_0^t \lambda(s)\mathrm{d}s}$
$f(t) =$	$F'(t)$	$-S'(t)$	$f(t)$	$\lambda(t)e^{-\int_0^t \lambda(s)\mathrm{d}s}$
$\lambda(t) =$	$\frac{F'(t)}{1 - F(t)}$	$-\frac{\mathrm{d}}{\mathrm{d}t}\ln S(t)$	$\frac{f(t)}{\int_t^\infty f(s)\mathrm{d}s}$	$\lambda(t)$

e.g. **Example 6.1.**

If $X \sim \mathrm{Exp}(\lambda)$, then

$$\lambda(t) = \frac{f(t)}{S(t)} = \frac{\lambda e^{-\lambda t}}{1 - (1 - e^{-\lambda t})} = \lambda.$$

Hence the exponential random variable has a constant hazard rate, i.e., old subject will be as likely to “die” as young subject, without regarding to their ages. Due to such memoryless property, the exponential distribution is generally not a reasonable model for the survival time of subjects with natural aging property.

e.g. **Example 6.2.**

Usually it would be more reasonable to model the lifetime of an item by an increasing hazard rate function rather than a constant hazard rate. ("Older" item will have higher chance to fail/die.) For example, we can use a linear hazard rate function $\lambda(t) = a + bt$. Then the distribution function is given by

$$\begin{aligned} F(x) &= 1 - \exp\left(-\int_0^x (a+bt)\mathrm{d}t\right) = 1 - \exp\left(-\left[at + \frac{bt^2}{2}\right]_0^x\right) \\ &= 1 - \exp\left(-ax - \frac{bx^2}{2}\right), \qquad \text{for } x > 0. \\ f(x) &= F'(x) = (a+bx)\exp\left(-ax - \frac{bx^2}{2}\right), \qquad \text{for } x > 0. \end{aligned}$$

In particular, if $a = 0$, the corresponding random variable is said to have the *Rayleigh distribution*.

---

§2 Indicator function

1、Definition: Indicator function (指示函数)

一个 indicator function of a subset A of a set Ω 为一个 function $\mathbb{1}_A : \Omega \to \{0,1\}$ 定义为

$$\mathbb{1}_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases}$$ (不表示概率分布)

2、Properties: indicator function 的性质

1° $\mathbb{1}_{A\cap B} = \min\{\mathbb{1}_A, \mathbb{1}_B\} = \mathbb{1}_A \times \mathbb{1}_B$

2° $\mathbb{1}_{A\cup B} = \max\{\mathbb{1}_A, \mathbb{1}_B\} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \times \mathbb{1}_B$

3° $\mathbb{1}_{A^c} = 1 - \mathbb{1}_A$

4° $E(\mathbb{1}_A) = Pr(A)$ i.e. $\int_{-\infty}^{\infty} \mathbb{1}_A f(x)\,dx = \int_A f(x)\,dx$

证明: 仅证明 1°

Consider using *truth table*...

For events, define ✓: occurs ×: does not occur

A	B	$A\cap B$	$\mathbb{1}_{A\cap B}$	$\mathbb{1}_A$	$\mathbb{1}_B$	$\min\{\mathbb{1}_A, \mathbb{1}_B\}$	$\mathbb{1}_A \times \mathbb{1}_B$
✓	✓	✓	1	1	1	$\min\{1,1\} = 1$	$1 \times 1 = 1$
✓	×	×	0	1	0	$\min\{1,0\} = 0$	$1 \times 0 = 0$
×	✓	×	0	0	1	$\min\{0,1\} = 0$	$0 \times 1 = 0$
×	×	×	0	0	0	$\min\{0,0\} = 0$	$0 \times 0 = 0$

all possible scenarios

identical

3. Property: $\mathbb{1}$ 服从伯努利分布

$\mathbb{1}$ 服从伯努利分布，成功率 $p = Pr(A)$，即 $\mathbb{1}_A \sim Ber(p)$

4. Property: indicator function 的应用

1° 可以将某些 density function 改写成一个 more compact way

e.g. $f(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, & \text{for } 0<x<1 \\ 0, & \text{otherwise} \end{cases}$

可以改写为 $\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \cdot \mathbb{1}_{\{x\in(0,1)\}}$

e.g. A r.v. X follows a Pareto distribution with p.d.f.

$$\begin{aligned}
f(x) &= \begin{cases} 3x^{-4}, & \text{for } x > 1; \\ 0, & \text{otherwise,} \end{cases} \\
&= \begin{cases} 3x^{-4}, & \text{for } x > 1; \\ 0, & \text{for } x \not> 1, \end{cases} \\
&= \begin{cases} 1 \times 3x^{-4}, & \text{for } x > 1; \\ 0 \times 3x^{-4}, & \text{for } x \not> 1, \end{cases} \\
&= 3x^{-4} \times \mathbf{1}_{\{x>1\}} \\
&= 3x^{-4} \times \mathbf{1}_{\{x\in(1,\infty)\}} \\
&= \underbrace{3x^{-4} \times \mathbf{1}_{(1,\infty)}(x)}_{\text{as a function of } x \text{ completely}} \qquad \text{for } x \in \mathbb{R}
\end{aligned}$$

more convenient to use in some applications with algebraic operations

2° 可以将某些 random variables 改写成一个 more compact way

e.g. $X(\omega) = \begin{cases} 2016, & \text{if } \omega \in A \\ 2601, & \text{if } \omega \in B \\ 6201, & \text{if } \omega \in C \end{cases}$

可以改写为 $X = 2016 \times \mathbb{1}_A + 2601 \times \mathbb{1}_B + 6210 \times \mathbb{1}_C$

或 $X = 2016^{\mathbb{1}_A} \times 2601^{\mathbb{1}_B} \times 6210^{\mathbb{1}_C}$

3° 利用 indicator function 求期望

e.g. **Example 6.3.**

A monkey types at a 26-letter keyboard with one key corresponding to each of the upper-case English letters. Each keystroke is chosen independently and uniformly at random from the 26 possibilities. If the monkey types 1 million letters, what is the expected number of times the sequence "BOOK" appears?

Solution:

There are $1,000,000 - 4 + 1 = 999,997$ places where the string "BOOK" can appear, each with a (non-independent) probability of $\frac{1}{26^4}$ of happening. If X is the random variable that counts the number of times the string "BOOK" appears, and X_i is the indicator variable that is 1 if the string "BOOK" appears starting at the i-th letter, then

$$\begin{aligned}
X &= X_1 + X_2 + \cdots + X_{999,997}, \\
\mathrm{E}(X) &= \mathrm{E}(X_1 + X_2 + \cdots + X_{999,997}) \\
&= \mathrm{E}(X_1) + \mathrm{E}(X_2) + \cdots + \mathrm{E}(X_{999,997}) \\
&= \frac{1}{26^4} + \frac{1}{26^4} + \cdots + \frac{1}{26^4} \\
&= \frac{999,997}{26^4} \\
&\approx 2.19.
\end{aligned}$$

★

e.g. **Example 6.4.**

A building has n floors numbered $1, 2, \ldots, n$, plus a ground floor. At the ground floor, m people get on the elevator together, and each gets off at a uniformly random one of the n floors (independently of everybody else). What is the expected number of floors the elevator stops at (not counting the ground floor)?

Solution:

Let X_i be the indicator that the elevator stopped at floor i.

$$\Pr(X_i = 1) = 1 - \Pr(\text{no one gets off at floor } i) = 1 - \left(\frac{n-1}{n}\right)^m.$$

Let X be the number of floors the elevator stops at, then the required expected number is

$$\begin{aligned}
\mathrm{E}(X) &= \mathrm{E}(X_1 + X_2 + \cdots + X_n) \\
&= \mathrm{E}(X_1) + \mathrm{E}(X_2) + \cdots + \mathrm{E}(X_n) \\
&= n\left[1 - \left(\frac{n-1}{n}\right)^m\right].
\end{aligned}$$

★

e.g. **Example 6.5.**

A coin with a probability p to get a head is flipped n times. A "run" is a maximal sequence of consecutive flips that are all the same. (Thus, for example, the sequence HTHHHHTTH with $n = 8$ has five runs.) Show that the expected number of runs is $1 + 2(n-1)p(1-p)$.

Solution:

Let X_i be the indicator for the event that a run starts at the i-th toss. Let $X = X_1 + X_2 + \cdots + X_n$ be the random variable for the number of runs total. Obviously, $\mathrm{E}(X_1) = 1$. For $i > 1$,

$$\begin{aligned}
\mathrm{E}(X_i) &= \Pr(X_i = 1) \\
&= \Pr\{i\text{-th toss is H}|(i-1)\text{-th toss is T}\} \times \Pr\{(i-1)\text{-th toss is T}\} \\
&\qquad + \Pr\{i\text{-th toss is T}|(i-1)\text{-th toss is H}\} \times \Pr\{(i-1)\text{-th toss is H}\} \\
&= p(1-p) + (1-p)p \\
&= 2p(1-p).
\end{aligned}$$

This gives

$$\begin{aligned}
\mathrm{E}(X) &= \mathrm{E}(X_1 + X_2 + \cdots + X_n) \\
&= \mathrm{E}(X_1) + \mathrm{E}(X_2) + \cdots + \mathrm{E}(X_n) \\
&= \mathrm{E}(X_1) + [\mathrm{E}(X_2) + \cdots + \mathrm{E}(X_n)] \\
&= 1 + (n-1) \times 2p(1-p) \\
&= 1 + 2(n-1)p(1-p).
\end{aligned}$$

★