Lecturer: Prof. Dr. Mete SONER
Coordinator: Yilin WANG

ETH Zürich

Solution Series 6

Q1. Suppose that internet users access a particular Web site according to a Poisson process with rate Λ per hour, but Λ is unknown. The Web site maintainer believes that Λ has a continuous distribution with probability density function:

$$f(\lambda) = \begin{cases} 2e^{-2\lambda} & \text{for } \lambda > 0 \\ 0 & \text{otherwise} \end{cases}.$$

Let X be the number of users who access the Web site during a one-hour period. If $X = 1$ is observed, find the conditional p.d.f. of Λ given $X = 1$.

Solution:

We compute first the CDF of λ given $X = 1$.

By hypothesis,

$$\mathbb{P}(X = 1 \,|\, \Lambda) = \mathbb{E}(\mathbf{1}_{X=1} \,|\, \Lambda) = e^{-\Lambda}\Lambda.$$

Hence

$$\begin{aligned} \mathbb{P}(X = 1) = \mathbb{E}(\mathbf{1}_{X=1}) = \mathbb{E}[\mathbb{E}(\mathbf{1}_{X=1} \,|\, \Lambda)] &= \int_0^\infty e^{-\lambda}\lambda f(\lambda) d\lambda \\ &= \int_0^\infty 2e^{-\lambda}\lambda e^{-2\lambda} d\lambda \\ &= \int_0^\infty 2\lambda e^{-3\lambda} d\lambda = 2/9. \end{aligned}$$

By definition of the conditional expectation, and that for any $x > 0$, $\mathbf{1}_{\Lambda\le\lambda}$ is $\sigma(\Lambda)$-measurable, hence

$$\mathbb{P}(X = 1, \Lambda \le \lambda) = \mathbb{E}[\mathbf{1}_{X=1}\mathbf{1}_{\Lambda\le\lambda}] = \mathbb{E}[\mathbb{E}(\mathbf{1}_{X=1} \,|\, \Lambda)\mathbf{1}_{\Lambda\le\lambda}] = \mathbb{E}[e^{-\Lambda}\Lambda\mathbf{1}_{\Lambda\le\lambda}],$$

and the CDF of λ given $X = 1$ is

$$\mathbb{P}(\Lambda \le \lambda \,|\, X = 1) = \frac{\mathbb{P}(X = 1, \Lambda \le \lambda)}{\mathbb{P}(X = 1)} = \frac{\int_0^\lambda 2se^{-3s} ds}{P(X = 1)}.$$

By differentiate the CDF w.r.t. λ, we obtain the PDF:

$$f(\lambda \,|\, X = 1) = 9\lambda e^{-3\lambda}.$$

Q2. Let X be a real-valued random variable. We define the characteristic function of X by

$$\begin{aligned}\varphi_X &: \mathbb{R} \to \mathbb{C} \\ t &\mapsto \varphi_X(t) := \mathbb{E}[e^{itX}] = \int e^{itx}\mu(dx),\end{aligned}$$

in which μ is the distribution of X on $\mathbb{R}$. It represents an important analytic tool, that the distribution of a random variable is uniquely determined (characterized) by the characteristic function. Show the following features:

(a)
- $\varphi_X(0) = 1$,
- $|\varphi_X(t)| \leq 1$,
- φ_X is continuous, and
- $\varphi_{aX+b}(t) = e^{itb}\varphi_X(at)$ for all $a, b \in \mathbb{R}$.

(b) Show that if the n-th moment of X exists, i.e. $\mathbb{E}[|X|^n] < \infty$, then φ_X is n times differentiable, and

$$\varphi_X^{(k)}(t) = i^k \mathbb{E}[X^k e^{itX}], \quad \text{for all } k \leq n,$$

(in particular $\varphi_X^{(k)}(0) = i^k \mathbb{E}[X^k]$).

Hint: one can use the inequality $|\frac{e^{i\alpha}-1}{\alpha}| \leq 1$, ($\alpha \in \mathbb{R}$).

(c) Compute the characteristic function for the standard normal distribution $\mathcal{N}(0,1)$, then for $\mathcal{N}(\mu, \sigma^2)$.

(d) Let X and Y be two independent random variables, defined on the same probability space. What is the characteristic function of $X + Y$?

Solution

(a)
- $\varphi_X(0) = \mathbb{E}[1] = 1$.
- $\varphi_X(t) = \int e^{itx}\mu(dx) \leq \int |e^{itx}|\mu(dx) = 1$, since μ is a probability measure on $\mathbb{R}$.
- This follows from the classical dominated convergence theorem and that e^{itx} is bounded.
- $\varphi_{aX+b}(t) = \mathbb{E}[e^{it(aX+b)}] = e^{itb}\mathbb{E}[e^{itaX}] = e^{itb}\varphi_X(at)$.

(b) We can prove it by induction. The case $k = 0$ is by definition of φ_X. Assume that for $k < n$ the proposition is true,

$$\frac{\varphi_X^{(k)}(t+h) - \varphi_X^{(k)}(t)}{h} = i^k \mathbb{E}[X^k e^{itX} X \frac{e^{ihX} - 1}{hX}],$$

since $|\frac{e^{ihX}-1}{hX}| \leq 1$ for any X and goes to i as $h \to 0$. By dominated convergence, using the fact that X^{k+1} is integrable, the above expression tends to $i^{k+1}\mathbb{E}[X^{k+1}e^{itX}]$.

(c) Let $X \sim \mathcal{N}(0,1)$,

$$\begin{aligned}\varphi_X(t) &= \int_{-\infty}^{+\infty} e^{itx}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx \\ &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}}e^{-(x^2-2itx)/2}dx \\ &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}}e^{-(x-it)^2/2}e^{-t^2/2}dx \\ &= e^{-t^2/2}.\end{aligned}$$

If $Y \sim \mathcal{N}(\mu,\sigma^2)$, then $(Y-\mu)/\sigma \sim \mathcal{N}(0,1)$. Thus

$$\varphi_Y(t) = e^{it\mu}\varphi_X(\sigma t) = e^{it\mu - \sigma^2 t^2/2}.$$

(d) As X and Y are independent, for every $f, g : \mathbb{R} \to \mathbb{R}$, $f(X)$ and $g(Y)$ are also independent random variables. Thus

$$\varphi_{X+Y}(t) = \mathbb{E}[e^{it(X+Y)}] = \mathbb{E}[e^{itX}e^{itY}] = \mathbb{E}[e^{itX}]\mathbb{E}[e^{itY}] = \varphi_X(t)\varphi_Y(t).$$

Let $X_1, X_2, \cdots, X_n$ be random variables defined on the same probability space. $(X_1, X_2, \cdots, X_n)$ is said to be a *Gaussian vector* if all $\mathbb{R}$-linear combinations of X_i are centered Gaussian random variable.

Q3. Let X and Y two independent standard normal random variables ($\mathcal{N}(0,1)$). Define the random variable

$$Z := \begin{cases} X & \text{if } Y \geq 0, \\ -X & \text{if } Y < 0. \end{cases}$$

(a) Compute the distribution of Z.

(b) Compute the correlation between X and Z.

(c) Compute $\mathbb{P}(X + Z = 0)$.

(d) Does (X, Z) follow a multivariate normal distribution (in other words, is (X, Z) a Gaussian vector)?

Solution:

(a) We just have to compute:

$$\begin{aligned}\mathbb{P}(Z \geq t) &= \mathbb{P}(\{X \geq t, Y > 0\} \cup \{X \leq -t, Y \leq 0\}) \\ &= \frac{1}{2}\mathbb{P}(X \geq t) + \frac{1}{2}\mathbb{P}(X \leq -t) \\ &= \mathbb{P}(X \geq t).\end{aligned}$$

So Z has the same law as X, thus $Z \sim \mathcal{N}(0,1)$.

(b) Using the definition of covariance

$$\begin{aligned} Cov(X,Z) &= \mathbb{E}(XZ) - \mathbb{E}(X)\mathbb{E}(Z) \\ &= \mathbb{E}\left(X^2 \mathbf{1}_{\{y \geq 0\}}\right) + \mathbb{E}\left(-X^2 \mathbf{1}_{\{y<0\}}\right) = 0. \end{aligned}$$

(c) We have that

$$\mathbb{P}(X+Z=0) = \mathbb{P}(Y<0) + \mathbb{P}(Y \geq 0, 2X=0) = \frac{1}{2}.$$

(d) It's not a multivariate normal, because the sum of them is not a normal.

Q4. Assume that $X := (X_1, X_2, \cdots, X_n)$ is a Gaussian vector, K_X the covariance matrix of X, which is defined by

$$K_X(i,j) = Cov(X_i, X_j).$$

(a) Let $\alpha_1, \cdots, \alpha_n$ be n real numbers, what is the law of $\sum_{i=1}^n \alpha_i X_i$ in terms of K_X?

(b) What can you say about K_X?

(c) If $K_X(1,2) = 0$, show that X_1 and X_2 are independent. Is that true if we don't assume that (X_1, X_2) is a Gaussian vector?

Hint: The characteristic function of the pair of random variables $X := (X_1, X_2)$, which is defined as

$$\begin{aligned} &\varphi_X : \mathbb{R}^2 \to \mathbb{C} \\ &t = (a,b) \mapsto \varphi_X(t) := \mathbb{E}[e^{it \cdot X}] = \mathbb{E}[e^{i(aX_1 + bX_2)}], \end{aligned}$$

characterizes also the joint law of (X_1, X_2).

Solution:

(a) As X is Gaussian vector, $Y := \sum_{i=1}^n \alpha_i X_i$ is also Gaussian. $\mathbb{E}(Y) = 0$ and

$$Var(Y) = Cov\left(\sum_{i=1}^n \alpha_i X_i, \sum_{j=1}^n \alpha_j X_j\right).$$

Cov is bilinear, hence

$$Var(Y) = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j K_X(i,j) = \alpha K_X {}^t\alpha,$$

where $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$.

(b) K_X is symmetric, and as Var is always non-negative, hence K_X is symmetric positive.

(c) First, $Z := (X_1, X_2)$ is Gaussian vector. The characteristic function of Z is

$$\begin{aligned}\varphi_Z(a,b) &= \mathbb{E}[e^{i(aX_1+bX_2)}] \\ &= \exp\left(-(a,b)\begin{pmatrix} Var(X_1) & 0 \\ 0 & Var(X_2)\end{pmatrix}\begin{pmatrix} a \\ b\end{pmatrix}/2\right) \\ &= \exp\left(-(a^2Var(X_1)+b^2Var(X_2))/2\right) \\ &= \exp\left(-a^2Var(X_1)/2\right)\exp\left(-b^2Var(X_2)/2\right),\end{aligned}$$

which follows from Q2.c).

The above expression is also the characteristic function of the joint law of two independent centered random gaussian variables, with variance respectively $Var(X_1)$ and $Var(X_2)$. Since the characteristic function determines the joint law, X_1 and X_2 are independent. If we omit the hypothesis that (X_1, X_2) is Gaussian vector, then 0 covariance does not imply the independence, as the example given in Q3.