Recitation 2 – Conditional Distribution, Moments

Problem 1:

Consider the random variable X that takes one of the two values $\{+1,-1\}$ with equal probabilities.

Given is the random variable Z, independent of X, with PDF $f_Z(z) = \frac{1}{2}e^{-|z|}$.

Let us define the random variable Y as follows: $Y = X + Z$.

1. Find $f_{Y|X}(y|x)$.
2. Find $f_Y(y)$.
3. Find $p_{X|Y}(x|y)$.

Solution:

1. Given the event $X = x$, Y is the random variable Z displaced.

$$f_{Z|X}(z|x) = f_{Y|X}(y|x)\frac{\partial(y|x)}{\partial(z|x)} = f_{Y|X}(y|x)$$

$$f_{Y|X}(y|x) = f_{Z|X}(z|x) = f_{Z|X}(y-x|x)$$

where we made use of the formula:

$$f_Y(y) = \sum_{i=1}^{n} \frac{f_Z(z_i)}{|g'(z_i)|}$$

Notice that the equation $y = g(z) = z + x$ has only one solution and the derivative in the denominator is equal to one.

Since X and Z are independent:

$$f_{Y|X}(y|x) = f_{Z|X}(y-x|x) = f_Z(y-x) = \tfrac{1}{2}e^{-|y-x|}$$

2. Let us use the Law of Total Probability:

$$f_Y(y) = \sum_x p_X(x) f_{Y|X}(y|x) = \tfrac{1}{2} f_{Y|X}(y|1) + \tfrac{1}{2} f_{Y|X}(y|-1) = \tfrac{1}{4}e^{-|y-1|} + \tfrac{1}{4}e^{-|y+1|}$$

<u>Another way:</u> Seeing as X and Z are independent, the PDF of Y is the convolution of the PDF of Z and the PDF of X:

$$f_Y(y) = f_X(y) * f_Z(y) = \left[\tfrac{1}{2}\delta(y-1) + \tfrac{1}{2}\delta(y+1)\right] * \left[\tfrac{1}{2}e^{-|y|}\right] = \tfrac{1}{4}e^{-|y-1|} + \tfrac{1}{4}e^{-|y+1|}$$

3. According to Bayes' theorem:

$$p_{X|Y}(x|y)=\frac{f_{Y|X}(y|x)p_X(x)}{f_Y(y)}=\begin{cases}\dfrac{f_{Y|X}(y|1)p_X(1)}{f_Y(y)} & x=1\\ \dfrac{f_{Y|X}(y|-1)p_X(-1)}{f_Y(y)} & x=-1\end{cases}$$

$$=\begin{cases}\dfrac{\frac{1}{2}e^{-|y-1|}\cdot\frac{1}{2}}{\frac{1}{4}e^{-|y-1|}+\frac{1}{4}e^{-|y+1|}} & x=1\\ \dfrac{\frac{1}{2}e^{-|y+1|}\cdot\frac{1}{2}}{\frac{1}{4}e^{-|y-1|}+\frac{1}{4}e^{-|y+1|}} & x=-1\end{cases}\qquad=\begin{cases}\dfrac{e^{-|y-1|}}{e^{-|y-1|}+e^{-|y+1|}} & x=1\\ \dfrac{e^{-|y+1|}}{e^{-|y-1|}+e^{-|y+1|}} & x=-1\end{cases}$$

Problem 2:

Find the expected value and variance of the following variables:

1. Laplace random variable: $f_X(x)=\frac{\lambda}{2}\cdot e^{-\lambda|x|}$

2. Gaussian random variable: $f_X(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

3. (**) Cauchy random variable: $f_X(x)=\frac{1}{\pi}\cdot\frac{1}{x^2+1}$

Solution:

1.

$$E(X)=\frac{\lambda}{2}\int_{-\infty}^{\infty}xe^{-\lambda|x|}dx=\frac{\lambda}{2}\left[\int_{-\infty}^{0}xe^{+\lambda x}dx+\int_{0}^{\infty}xe^{-\lambda x}dx\right]$$

Let us make a change of variables in the first integral: $x\to -x$. We get:

$$E(X)=\frac{\lambda}{2}\left[-\int_{0}^{\infty}xe^{-\lambda x}dx+\int_{0}^{\infty}xe^{-\lambda x}dx\right]=\frac{\lambda}{2}\left[\frac{1}{\lambda^2}-\frac{1}{\lambda^2}\right]=0$$

Notice that, in general, it is not correct to say that the integrals cancel each other out. It is possible that the boundaries of the integrals go to infinity at a different rate, resulting in that the integrals take different values. If the integral converges absolutely, then this cannot be.

Now, let us find the second moment:

$$E(X^2)=\frac{\lambda}{2}\int_{-\infty}^{\infty}x^2e^{-\lambda|x|}dx=\frac{\lambda}{2}\left[\int_{-\infty}^{0}x^2e^{+\lambda x}dx+\int_{0}^{\infty}x^2e^{-\lambda x}dx\right]=\lambda\int_{0}^{\infty}x^2e^{-\lambda x}dx$$

We will use the following identity: (It can be proven using integration by parts)

$$\int_0^\infty x^k e^{-\lambda x} dx = \frac{k!}{\lambda^{k+1}} \qquad k = 0,1,2..... \ \lambda > 0$$

This results in:

$$E(X^2) = \frac{2}{\lambda^2}$$

$$\Rightarrow Var(X) = E(X^2) - [E(X)]^2 = \frac{2}{\lambda^2}$$

2. It can be calculated directly by solving the integral:

$$E(X) = \int_{-\infty}^{-\infty} x f_X(x) dx$$

However, a simpler way is to calculate it using its characteristic function.

<u>Reminder:</u> for a Gaussian variable $X \sim N(\mu, \sigma^2)$, the characteristic function is:

$$\varphi_X(\omega) = E(e^{i\omega X}) = e^{i\mu\omega - \frac{1}{2}\sigma^2\omega^2}$$

After deriving the above function twice with respect to $X \sim N(\mu, \sigma^2)$:

$$\frac{d}{d\omega}\varphi_X(\omega) = E(iXe^{i\omega X}) = (i\mu - \sigma^2\omega)e^{i\mu\omega - \frac{1}{2}\sigma^2\omega^2}$$
$$\frac{d^2}{d\omega^2}\varphi_X(\omega) = E(-X^2 e^{i\omega X}) = (i\mu - \sigma^2\omega)^2 e^{i\mu\omega - \frac{1}{2}\sigma^2\omega^2} - \sigma^2 e^{i\mu\omega - \frac{1}{2}\sigma^2\omega^2}$$

By replacing $\omega = 0$, we get:

$$\left.\frac{d}{d\omega}\phi_X(\omega)\right|_{\omega=0} = iE(X) = i\mu \Rightarrow E(X) = \mu$$
$$\left.\frac{d^2}{d\omega^2}\phi_X(\omega)\right|_{\omega=0} = -E(X^2) = -\mu^2 - \sigma^2 \Rightarrow E(X^2) = \mu^2 + \sigma^2 \Rightarrow Var(X) = \sigma^2$$

3. (**) Since the PDF "behaves" like $\frac{1}{x^2}$ at infinity, the following integral diverges:

$$\int_0^\infty x \cdot f_X(x) dx = \frac{1}{\pi}\int_0^\infty \frac{x}{x^2+1} dx \geq \frac{1}{\pi}\int_1^\infty \frac{x}{x^2+x^2} dx = \frac{1}{2\pi}\int_1^\infty \frac{1}{x} dx = \infty$$

Likewise, $\int_{-\infty}^{0} x \cdot f_X(x) dx = -\infty$

Thus, the expected value of a Cauchy random variable is **undefined** (and, thus, the variance is also undefined). Equivalently, the integral $\int_{-\infty}^{\infty} f_X(x)|x|dx$ diverges, hence the expected value is undefined.

<u>Note</u>: Seemingly, one could have claimed that the PDF is an even function and, thus, $x \cdot f_X(x)$ is odd, and from here it holds that:

$$\int_{-\infty}^{\infty} f_X(x)xdx = \lim_{M\to\infty} \int_{-M}^{M} f_X(x)xdx = \lim_{M\to\infty} 0 = 0$$

like in section 1. However, this claim is incorrect: one can use the identity $\int_{-\infty}^{\infty} f_X(x)xdx = \lim_{M\to\infty} \int_{-M}^{M} f_X(x)xdx$ only when the integral **converges absolutely.** Otherwise, the integral $\int_{-\infty}^{\infty} x \cdot f_X(x)dx$ is **undefined**.

<u>Problem 3:</u>

Consider the Gaussian random variable $X \sim N(0,\sigma^2)$. Prove that:

$$\mathrm{E}\left[X^n\right] = \begin{cases} 0, & n \quad odd \\ \sigma^n \cdot 1\cdot 3\cdot 5\cdot \ldots \cdot (n-1), & n \quad even \end{cases}$$

<u>Solution:</u>

If n is odd, then:

$$\mathrm{E}\left[X^n\right] = \int_{-\infty}^{\infty} x^n f_X(x)dx = 0$$

If n is even, we use the characteristic function of X:

$$\Phi_X(\omega) = \mathrm{E}\left[e^{j\omega X}\right] = e^{-\frac{1}{2}\omega^2\sigma^2} = \sum_{k=0}^{\infty}\left(-\frac{1}{2}\sigma^2\right)^k \frac{1}{k!}\omega^{2k}$$

$$\mathrm{E}\left[X^{2m}\right] = \frac{1}{j^{2m}} \frac{d^{2m}\Phi_X(\omega)}{d\omega^{2m}}\bigg|_{\omega=0} = \frac{1}{(-1)^m}\left(-\frac{1}{2}\sigma^2\right)^m \frac{(2m)!}{m!} = \sigma^{2m}\cdot(2m-1)\cdot(2m-3)\cdot\ldots\cdot 1$$

Extra Questions

Problem 4:

1. Consider the continuous random variable X with CDF $F_X(x)$.
Assume $f_X(x) > 0$ for all real x.
Let us define the random variable $Z = F_X(X)$. Prove that Z is uniformly distributed in the range [0,1].

2. Consider the random variable U, which uniformly distributes in the range [0,1].
We would like to find a random variable Y with a monotone, strictly increasing and continuous CDF $F_Y(y)$.
Let us denote the inverse function of $F_Y(y)$ by $F_Y^{-1}(y)$.
Prove that the CDF of random variable $\tilde{Y} = F_Y^{-1}(U)$ is $F_Y(y)$.

Solution:

1. Let us solve in two ways:

I) Solution based on definitions. First, we will find the CDF of Z.

Notice that $0 \le F_X(x) \le 1$ for all x, it holds that:

$$F_Z(z) = \Pr\{F_X(X) \le z\} \quad \begin{aligned} &= 0 \qquad \textit{when } z \le 0 \\ &= 1 \qquad \textit{when } z \ge 1 \end{aligned}$$

Now, for $0 < z < 1$:

$$F_Z(z) = \Pr\{F_X(X) \le z\} \underset{(1)}{=} \Pr\{X \le F_X^{-1}(z)\} = F_X(F_X^{-1}(z)) = z$$

(1) $F_X(\bullet)$ is monotonous and strictly increasing, and, thus, injective (has an inverse function).

In conclusion:

$$F_Z(z) = \begin{cases} 0 & z \le 0 \\ z & 0 < z < 1 \\ 1 & z \ge 1 \end{cases}$$

That is to say, $Z \sim Unif(0,1)$.

II) Solution based on the preposition.

For $z \le 0$ or $z \ge 1$, the equation $F_X(x) = z$ has no solutions and, therefore, $f_Z(z) = 0$.

Otherwise, there exists only one solution $x_1 = F_X^{-1}(z)$.

$$g(x) = F_X(x)$$
$$g'(x) = f_X(x) \Rightarrow g'(x_1) = f_X\left(F_X^{-1}(z)\right)$$

Thus, based on the preposition, we get:

$$f_Z(z) = \frac{f_X(x_1)}{|g'(x_1)|} = \frac{f_X\left(F_X^{-1}(z)\right)}{f_X\left(F_X^{-1}(z)\right)} = 1$$

Note: This claim holds even if $f_X(x) > 0$ is not true for all x.

2. Let us find the CDF of $\tilde{Y}$:

$$F_{\tilde{Y}}(\tilde{y}) = \Pr\{\tilde{Y} \le \tilde{y}\} = \Pr\{F_Y^{-1}(U) \le \tilde{y}\} = \Pr\{U \le F_Y(\tilde{y})\} \underset{0 \le F_Y(\tilde{y}) \le 1}{=} F_Y(\tilde{y})$$

<u>Appendix 1 - Moments</u>

<u>Definitions:</u>

Let X be a random variable. We shall define for it the following:

- Expected value: $E\{X\} = \int_{-\infty}^{\infty} \alpha \cdot f_X(\alpha) d\alpha$; it is usually symbolized as η or η_X.
- Variance: $Var\{X\} = \int_{-\infty}^{\infty} (\alpha - E\{X\})^2 \cdot f_X(\alpha) d\alpha = E\{X^2\} - E^2\{X\}$; it is usually symbolized as σ^2 or σ_X^2.
- Moment of order n: $m_n = E\{X^n\} = \int_{-\infty}^{\infty} \alpha^n \cdot f_X(\alpha) d\alpha$.
- Central moment of order n: $\mu_n = E\{(X - E\{X\})^n\}$.
  For a random variable with expected value 0, the moment of order n coincides with the central moment of order n.

If X is a discrete random variable, then, additionally:

- Expected value: $E\{X\} = \sum_i x_i \cdot \Pr(X = x_i)$.

- Variance: $Var\{X\} = \sum_i (x_i - E\{X\})^2 \cdot \Pr(X = x_i) = \sum_i x_i^2 \cdot \Pr(X = x_i) - E^2\{X\}$.
- Moment of order n: $m_n = E\{X^n\} = \sum_i x_i^n \Pr(X = x_i)$.

Characteristics:

- The expected value of a constant is the constant itself $E\{c\} = c$.
- Linearity: $E\{aX + b + cY\} = aE\{X\} + b + cE\{Y\}$ (where a, b, c are constants).
- For a deterministic function $g(\bullet)$, the expected value is:

$$E\{g(X)\} = \int_{-\infty}^{\infty} g(\alpha) \cdot f_X(\alpha) d\alpha$$

For a discrete random variable, one can also formulate:

$$E\{g(X)\} = \sum_i g(x_i) \Pr(X = x_i)$$

Notice:

- The expected value is the first moment, i.e. $E\{X\} = m_1$, and the variance is the second central moment: $Var\{X\} = \mu_2$.
- The variance of a random variable (whether discrete or continuous) is non-negative, meaning $Var\{X\} \geq 0$. When the variance is zero, it can be assumed that the variable is deterministic and equal to its expected value (more specifically, X is a constant in the sense of squared expectation). Mathematically: $Var\{X\} = 0 \Rightarrow X = \eta_X$.

  This characteristic can be seen in Gaussian random variables, for which, assuming variance zero, its PDF converts into a delta function at the expected value – the PDF of a deterministic variable.
- The expectation is defined when the integral (in the continuous case) or the sum (in the discrete case) converge absolutely. For example, for the continuous case we get: $\int_{-\infty}^{\infty} |\alpha| \cdot f_X(\alpha) d\alpha < \infty$.
- Calculation of the variance using moments: $Var\{X\} = m_2 - m_1^2$.

Appendix 2 – First Characteristic Function of Random Variables

For a random variable X, its first characteristic function (represented as $\phi_X(\omega)$) is defined as:

$$\phi_X(\omega) \equiv E\{e^{j\omega X}\} = \int_{-\infty}^{\infty} f_X(x) e^{j\omega x} dx$$

and the following holds:

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_X(\omega) e^{-j\omega x} d\omega$$

The characteristic function and the PDF are a Fourier pair and, thus, a one-to-one relationship exists between them[1].

One can obtain the moments of a random variable from the characteristic function as seen below:

$$m_n = E\{X^n\} = \frac{1}{j^n} \cdot \left. \frac{d^n \phi_X(\omega)}{d\omega^n} \right|_{\omega=0}$$

Moreover, if the function is analytical, we get the relationship (which is derived from an expansion of the Maclaurin series of $e^{j\omega X}$):

$$\phi_X(\omega) = \sum_{n=0}^{\infty} m_n \cdot \frac{(j\omega)^n}{n!}$$

(*) Conclusion: If the characteristic function is analytical, the sequence of moments of the random variable defines it uniquely, and, thus, defines uniquely the PDF. As seen in the lecture, generally speaking, the sequence of moments of a random variable does not define uniquely the distribution of the variable (this problem is named the **moment problem**).

Note:

If X is a discrete random variable, one may also write:

$$\phi_X(\omega) \equiv E\{e^{j\omega X}\} = \sum_i e^{j\omega x_i} \Pr\{X = x_i\}$$

1 Notice that we define here the Fourier transform of function $g(x)$ as follows:

$$G(\omega) = \int_{-\infty}^{\infty} g(t) e^{j\omega t} dt$$

This definition is different than the definition provided in the signal processing courses:

$$G_2(\omega) = \int_{-\infty}^{\infty} g(t) e^{-j\omega t} dt$$