ECE 361 Probability for Engineers (Fall, 2016)
Lecture 6b

---

§3.3 Normal RVs

The Gaussian or normal RV X has support $\mathcal{X} = \mathbb{R}$ and PDF

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}, \ x \in \mathbb{R}. \tag{1}$$

We write $X \sim N(\mu, \sigma)$ to denote that X is an RV that is normally distributed with parameters $\mu \in \mathbb{R}$ and $\sigma \geq 0$. Note: it is also common to write $X \sim N(\mu, \sigma^2)$. To see the import of this ambiguity, we interpret $X \sim N(0, 2)$ to mean $\sigma = 2$ but under the alternate notation we interpret $\sigma^2 = 2$. One can show (although it is not straightforward) that this is a valid PDF:

$$\int_{-\infty}^{\infty} f_X(x)\mathrm{d}x = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x = 1. \tag{2}$$

The CDF cannot be put in closed form, it can only be expressed (consistent with the definition of the CDF) as the integral of the PDF:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\mathrm{d}t = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} \mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}} \mathrm{d}t. \tag{3}$$

We call μ the mean parameter and σ^2 the variance parameter because:

$$\mathbb{E}[X] = \mu, \ \mathrm{var}(X) = \sigma^2. \tag{4}$$

Note the PDF is symmetric around μ, i.e., $f_X(x) = f_X(2\mu - x)$.

Claim: for any RV X (not just a normal RV) symmetric around some μ we have $\mathbb{P}(X > \mu) = \mathbb{P}(X \leq \mu) = 1/2$ and $\mathbb{E}[X] = \mu$.

Proof: First break the support into $(-\infty, \mu) \cup [\mu, \infty)$, then define $y = 2\mu - x$ and simplify:

$$\begin{aligned} 1 &= \int_{-\infty}^{\mu} f_X(x)\mathrm{d}x + \int_{\mu}^{\infty} f_X(x)\mathrm{d}x \\ &= \int_{-\infty}^{\mu} f_X(2\mu - x)\mathrm{d}x + \mathbb{P}(X > \mu) \\ &= \int_{\mu}^{\infty} f_X(y)\mathrm{d}y + \mathbb{P}(X > \mu) \\ &= 2\mathbb{P}(X > \mu). \end{aligned} \tag{5}$$

Next:

$$\begin{aligned} \mathbb{E}[X] &= \int_{-\infty}^{\mu} x f_X(x)\mathrm{d}x + \int_{\mu}^{\infty} x f_X(x)\mathrm{d}x \\ &= \int_{\mu}^{\infty} (2\mu - y) f_X(y)\mathrm{d}y + \int_{\mu}^{\infty} x f_X(x)\mathrm{d}x \\ &= 2\mu \int_{\mu}^{\infty} f_X(y)\mathrm{d}y = 2\mu\mathbb{P}(X > \mu) = \mu. \end{aligned} \tag{6}$$

Returning to the normal RV, we now show $\mathrm{var}(X) = \sigma^2$, use the definition and the change of variable $y = (x-\mu)/\sigma$:

$$\begin{aligned} \mathrm{var}(X) &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(x-\mu)^2 \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x \\ &= \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y^2 \mathrm{e}^{-\frac{y^2}{2}}\mathrm{d}y \\ &= \frac{\sigma^2}{\sqrt{2\pi}}\left(-y\mathrm{e}^{-\frac{y^2}{2}}\right)\Big|_{-\infty}^{\infty} + \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{y^2}{2}}\mathrm{d}y \\ &= \sigma^2. \end{aligned} \tag{7}$$

A key property of the normal distribution is that a linear function of a normal RV is a normal RV:

If $X \sim N(\mu,\sigma)$ is a normal RV then $Y = aX+b$ is a normal RV. Naturally $\mathbb{E}[Y] = a\mu + b$ and $\mathrm{var}(Y) = a^2\sigma^2$.

The standard normal RV

A *standard* normal RV is a normal RV with $\mu = 0$ and $\sigma = 1$ (zero mean and unit variance). We use the special symbol Φ for the CDF of a standard normal:

$$\Phi(y) = \mathbb{P}(Y \leq y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t. \tag{8}$$

Some values of the function $\Phi(y)$ are recorded in the table on page 155. Note the table only gives values for positive values y since negative values hold by symmetry: $\Phi(-y) = 1 - \Phi(y)$ for all y.

We *standardize* any RV (not just a normal by subtracting its mean and dividing by its standard deviation: if X is an RV with $\mathbb{E}[X] = \mu$ and $\mathrm{var}(X) = \sigma^2$ then the standardization of X is the RV $Y = (X-\mu)/\sigma$. Note:

$$\mathbb{E}[Y] = \mathbb{E}\left[\frac{X-\mu}{\sigma}\right] = 0, \ \mathrm{var}(Y) = \mathrm{var}\left(\frac{X-\mu}{\sigma}\right) = 1. \tag{9}$$

This standardization is particularly important for normal RVs: given an arbitrary normal RV $X \sim N(\mu,\sigma)$ we standardize to $Y = (X-\mu)/\sigma$ so that $Y \sim N(0,1)$. To be clear: one can always standardize an RV, but the claim that the standardized RV is of the same distribution as the original is special and does not hold in general (but it does hold for the normal).

Example. The annual snowfall is $X \sim N(\mu,\sigma)$ with $\mu = 60$ inches and $\sigma = 20$ inches. Find the probability the snowfall will be at least $x = 80$ inches. Note $(x-\mu)/\sigma = 1$.

$$\mathbb{P}(X > x) = \mathbb{P}\left(\frac{X-\mu}{\sigma} > \frac{x-\mu}{\sigma}\right) = \mathbb{P}(Y > 1) = 1 - \mathbb{P}(Y \leq 1) = 1 - \Phi(1) \approx 1 - 0.8413 = 0.1587. \tag{10}$$

Thus this example illustrates that for $X \sim N(\mu,\sigma)$ we have

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = \mathbb{P}\left(Y \leq \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right). \tag{11}$$

§3.4 Joint PDFs of multiple RVs

Define a pair of RVs (X,Y) as jointly continuous if there exists a function $f_{X,Y}(x,y)$ (the joint PDF) such that

$$\mathbb{P}((X,Y) \in B) = \int_{(x,y)\in B} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y \tag{12}$$

for all $B \subset \mathbb{R}^2$. Letting $B = \mathbb{R}^2$ we see the joint PDF must obey the normalization property:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y = 1. \tag{13}$$

Consider some point (a, c) and some small δ and form the set $B = [a, a+\delta] \times [c, c+\delta]$ with area δ^2. Then

$$\mathbb{P}((X,Y) \in B) = \int_c^{c+\delta} \int_a^{a+\delta} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y = f_{X,Y}(a,c)\delta^2, \tag{14}$$

so $f_{X,Y}(x,y)$ is the "probability per unit area" at point (x,y).

Note we can always obtain the marginal PDFs from the joint. Indeed for any $A \subset \mathbb{R}$ we have

$$\mathbb{P}(X \in A) = \mathbb{P}(X \in A \cap Y \in \mathbb{R}) = \int_A \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}y\mathrm{d}x. \tag{15}$$

Thus the marginal PDFs are:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}y, \;\; f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}x. \tag{16}$$

Example. (2D uniform PDF.) Romeo and Juliet each arrive at their meeting place at independent random times selected uniformly over $[0,1]$. Thus $f_{X,Y}(x,y) = c$ for some c for $(x,y) \in [0,1]^2$ and 0 else. By integrating over $[0,1]^2$ we see:

$$1 = \int_{[0,1]^2} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y = c. \tag{17}$$

More generally, if (X,Y) are independent and uniformly distributed over a set $S \subset \mathbb{R}^2$ and we are interested in $(X,Y) \in A$ for some $A \subset S$, then

$$\mathbb{P}((X,Y) \in A) = \frac{\mathrm{Area}(A)}{\mathrm{Area}(S)}. \tag{18}$$

Example. See Fig. 3.12. Suppose (X,Y) are independent and uniform over the set S in Fig. 3.12. Find the marginals for X, Y.

Example. (Buffon's needle.) See Fig. 3.13. A surface is ruled with parallel lines separated by distance d. We drop a needle of length $l < d$. What is the probability the needle lies across one of the lines? Let (X, Θ) be a pair of RVs as in Fig. 3.13 where $x \in [0, d/2]$ and $\theta \in [0, \pi/2]$ with area $\pi d/4$. As both X, Θ are uniform it follows that $f_{X,\Theta}(x,\theta) = 4/(\pi d)$ for (x,θ) in that set. The event the needle hits the line is the event $X \leq \frac{l}{2}\sin\Theta$ and thus:

$$\mathbb{P}(X \leq \frac{l}{2}\sin\Theta) = \int_{(x,\theta): x \leq \frac{l}{2}\sin\theta} f_{X,\Theta}(x,\theta)\mathrm{d}x\mathrm{d}\theta = \frac{4}{\pi d}\int_0^{\pi/2}\int_0^{(l/2)\sin\theta} \mathrm{d}x\mathrm{d}\theta = \cdots = \frac{2l}{\pi d}. \tag{19}$$

Note this experiment is easily performed, and we can approximate the probability by estimating the fraction of dropped needles that intersect a line – this approximation becomes exact as the number of independent trials grows to infinity. Suppose that our approximation is $\hat{p}$, then we can find:

$$\hat{\pi} = \frac{2l}{\hat{p}d} \approx \pi. \tag{20}$$

In other words, we can estimate the value of π by using $\hat{\pi}$ which is a function of the constants l, d and our approximation $\hat{p}$.

Joint CDFs

The joint CDF of two RVs (X,Y) is

$$F_{X,Y}(x,y) = \mathbb{P}(X \leq x, Y \leq y). \tag{21}$$

The joint CDF can be found from the joint PDF by integration:

$$F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(s,t)\mathrm{d}t\mathrm{d}s. \tag{22}$$

The joint PDF can be found from the joint CDF by differentiation:

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y). \tag{23}$$

Example. Let (X,Y) be independent uniform RVs over the unit square. The joint CDF is:

$$F_{X,Y}(x,y) = \mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y) = xy, \ (x,y) \in [0,1]^2. \tag{24}$$

The joint PDF is:

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} xy = 1. \tag{25}$$

Expectation

Let $Z = g(X,Y)$ be a function of the pair of RVs (X,Y). Then:

$$\mathbb{E}[Z] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y) \mathrm{d}x \mathrm{d}y. \tag{26}$$

Further if $Z = aX + bY + c$ then $\mathbb{E}[Z] = a\mathbb{E}[X] + b\mathbb{E}[Y] + c$.

More than two RVs

In perfect analogy with the case of two RVs, we have for three RVs (X,Y,Z):

$$\mathbb{P}((X,Y,Z) \in B) = \int_{(x,y,z) \in B} f_{X,Y,Z}(x,y,z) \mathrm{d}x \mathrm{d}y \mathrm{d}z, \ \forall B \subset \mathbb{R}^3. \tag{27}$$

We can marginalize Z by integrating over it to find the joint for (X,Y):

$$f_{X,Y}(x,y) = \int_{-\infty}^{\infty} f_{X,Y,Z}(x,y,z) \mathrm{d}z, \tag{28}$$

and we can marginalize (Y,Z) by integrating over them to find the marginal for X:

$$f_X(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y,Z}(x,y,z) \mathrm{d}y \mathrm{d}z. \tag{29}$$

Naturally the expecation for $W = g(X,Y,Z)$ is:

$$\mathbb{E}[W] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y,z) f_{X,Y,Z}(x,y,z) \mathrm{d}x \mathrm{d}y \mathrm{d}z. \tag{30}$$

and naturally for $W = aX + bY + cZ$ we have $\mathbb{E}[W] = a\mathbb{E}[X] + b\mathbb{E}[Y] + c\mathbb{E}[Z]$. More generally, for RVs $(X_1, \ldots, X_n)$ and scalars $(a_1, \ldots, a_n)$ we have:

$$\mathbb{E}[a_1 X_1 + \cdots a_n X_n] = \mathbb{E}\left[\sum_{i=1}^n a_i X_i\right] = \sum_{i=1}^n a_i \mathbb{E}[X_i] = a_1 \mathbb{E}[X_1] + \cdots + a_n \mathbb{E}[X_n]. \tag{31}$$

§3.5 Conditioning

Conditioning an RV on an event

Given an event A with $\mathbb{P}(A) > 0$ the conditional PDF $f_{X|A}(x)$ is defined as the function for which:

$$\mathbb{P}(X \in B | A) = \int_B f_{X|A}(x) \mathrm{d}x, \ \forall B \subset \mathbb{R}. \tag{32}$$

Again, by choosing $B = \mathbb{R}$ we require normalization:

$$1 = \int_{-\infty}^{\infty} f_{X|A}(x) \mathrm{d}x. \tag{33}$$

For events of the form $\{X \in A\}$ we find

$$\mathbb{P}(X \in B|X \in A) = \frac{\mathbb{P}(X \in B, X \in A)}{\mathbb{P}(X \in A)} = \frac{1}{\mathbb{P}(X \in A)} \int_{A \cap B} f_X(x) \mathrm{d}x \tag{34}$$

which means the conditional PDF is

$$f_{X|\{X \in A\}}(x) = \begin{cases} \frac{f_X(x)}{\mathbb{P}(X \in A)}, & x \in A \\ 0, & \text{else} \end{cases} \tag{35}$$

Example. (Exponential RV is memoryless.) Suppose the lifetime of a lightbulb is an exponential RV with parameter λ, i.e., $T \sim \mathrm{Exp}(\lambda)$. Given $T > t$, find the distribution for the additional lifetime X of the lightbulb. Let $A = \{T > t\}$. Then:

$$\mathbb{P}(X > x|A) = \mathbb{P}(T > t + x|T > t) = \frac{\mathbb{P}(T > t + x, T > t)}{\mathbb{P}(T > t)} = \frac{\mathbb{P}(T > t + x)}{\mathbb{P}(T > t)} = \frac{\mathrm{e}^{-\lambda(t+x)}}{\mathrm{e}^{-\lambda t}} = \mathrm{e}^{-\lambda x}. \tag{36}$$

In other words, $\mathbb{P}(X > x|A) = \mathbb{P}(X > x)$, i.e., the additional lifetime of the lightbulb is independent of the past lifetime. This is the memorylessness property of the exponential distribution.

References

[1] *Introduction to Probability, 2nd Edition* by Dimitri P. Bertsekas and John N. Tsitsiklis, Athina Scientific Press, 2008.