UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Homework 4
Spring 2024

1. **Change of Variables**

   Let X be a continuous random variable with cdf F_X and pdf $f_X > 0$ everywhere, and let $Y = g(X)$, where g is a differentiable function.

   a. Suppose that g is also invertible. Find the pdf of Y, f_Y, in terms of g and f_X.
   b. Let $U \sim \text{Uniform}([0,1])$. Using the conclusion from part a, show that $F_X^{-1}(U)$ has the same distribution as X. (This allows us to generate a given random variable given only a uniform random number generator.)
   c. Now suppose that $g(x) = x^2$. Find the pdf of Y in terms of the pdf of X. Also find the pdf of Y when X is a standard normal random variable in particular.
      (Note that this g is not invertible, unlike in part a.)

   Solution:

   a. g is a continuous invertible function from $\mathbb{R}$ to $\mathbb{R}$, so g must be monotonic, i.e. strictly increasing or strictly decreasing. Let us first find the cdf of Y:

   $$F_Y(y) = \mathbb{P}(g(X) \le y) = \begin{cases} \mathbb{P}(X \le g^{-1}(y)) = F_X(g^{-1}(y)) & \text{if } g \text{ is increasing} \\ \mathbb{P}(X \ge g^{-1}(y)) = 1 - F_X(g^{-1}(y)) & \text{if } g \text{ is decreasing.} \end{cases}$$

   Then, by the chain rule of differentiation, we find the pdf of Y as

   $$f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} F_Y(y) = \left| \frac{\mathrm{d}}{\mathrm{d}y} F_X(g^{-1}(y)) \right| = f_X(g^{-1}(y)) \cdot \left| \frac{\mathrm{d}}{\mathrm{d}y} g^{-1}(y) \right|.$$

   Using the inverse function rule, we can further simplify to

   $$f_Y(y) = f_X(g^{-1}(y)) \cdot \frac{1}{|g'(g^{-1}(y))|}.$$

   b. Let $Y = F_X^{-1}(U)$. F_X is differentiable because X is a continuous random variable, and strictly increasing because $f_X > 0$ everywhere, so its inverse F_X^{-1} is also differentiable and monotonically increasing. Using the conclusion of part a with $g = F_X^{-1}$,

   $$F_Y(y) = F_U(g^{-1}(y)) = F_U(F_X(y)) = F_X(y),$$

   which shows that Y has the same distribution as X. Note that $F_U(u) = \mathbb{P}(U \le u) = u$ for $U \sim \text{Uniform}([0,1])$.

c. The cdf of $Y = X^2$ is

$$\mathbb{P}(X^2 \leq y) = \mathbb{P}(-\sqrt{y} \leq X \leq \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)\,\mathrm{d}x.$$

By the fundamental theorem of calculus, the pdf of Y is

$$f_Y(y) = \frac{1}{2\sqrt{y}}(f_X(-\sqrt{y}) + f_X(\sqrt{y})).$$

For $X \sim \mathcal{N}(0,1)$, the pdf of X^2 evaluates to

$$f_Y(y) = \frac{1}{2\sqrt{y}}\left(\frac{1}{\sqrt{2\pi}}e^{-y/2} + \frac{1}{\sqrt{2\pi}}e^{-y/2}\right) = \frac{1}{\sqrt{2\pi y}}e^{-y/2}.$$

(This is known as the **chi-squared** distribution with 1 degree of freedom.)

2. **Convolution practice**

The random variables X, Y, and Z are independent and uniformly distributed between zero and one. Find the PDF of X + Y + Z.

Solution: Let $V = X + Y$. As in Example 4.10, the PDF of V is

$$f_V(v) = \begin{cases} v, & 0 \leq v \leq 1, \\ 2 - v, & 1 \leq v \leq 2, \\ 0, & \text{otherwise.} \end{cases}$$

Let $W = X + Y + Z = V + Z$. We convolve the PDFs f_V and f_Z, to obtain

$$f_W(w) = \int f_V(v) f_Z(w - v)\, dv.$$

We first need to determine the limits of the integration. Since $f_V(v) = 0$ outside the range $0 \leq v \leq 2$, and $f_W(w - v) = 0$ outside the range $0 \leq w - v \leq 1$, we see that the integrand can be nonzero only if

$$0 \leq v \leq 2, \quad \text{and} \quad w - 1 \leq v \leq w.$$

We consider three separate cases. If $w \leq 1$, we have

$$f_W(w) = \int_0^w v f_V(v) f_Z(w - v)\, dv = \frac{w^2}{2}.$$

If $1 \leq w \leq 2$, we have

$$\begin{aligned} f_W(w) &= \int_{w-1}^1 v f_V(v) f_Z(w - v)\, dv + \int_1^w (2 - v) f_V(v) f_Z(w - v)\, dv \\ &= \frac{1}{2}\left((w - 1)^2 - \frac{(w - 2)^2}{2} \right) + \frac{1}{2}. \end{aligned}$$

Finally, if $2 \leq w \leq 3$, we have

$$f_W(w) = \int_{w-1}^2 (2 - v) f_V(v) f_Z(w - v)\, dv = \frac{(3 - w)^2}{2}.$$

To summarize,

$$f_W(w) = \begin{cases} \frac{w^2}{2}, & 0 \leq w \leq 1, \\ 1 - \frac{(w-1)^2}{2} - \frac{(2-w)^2}{2}, & 1 \leq w \leq 2, \\ \frac{(3-w)^2}{2}, & 2 \leq w \leq 3, \\ 0, & \text{otherwise.} \end{cases}$$

3. **Moment-Generating Functions Practice**

The **moment-generating function** (mgf) of a random variable X is the function

$$M_X(s) = \mathbb{E}(e^{sX}) = \mathbb{E}\left(\sum_{k=0}^{\infty} \frac{(sX)^k}{k!}\right) = \sum_{k=0}^{\infty} \frac{s^k}{k!}\,\mathbb{E}(X^k).$$

In this class, we will not worry about technical details about the convergence of Taylor series, so we will say that the mgf is equal to any of the expressions above.

The mgf gets its name because it is the function *generating* the moments $\mathbb{E}(X^p)$, $p \geq 1$, of X. More specifically, by evaluating the pth derivative of the mgf at $s = 0$, we have a method to explicitly find the pth moment of X from its mgf:

$$\left[\frac{\mathrm{d}^p}{\mathrm{d}s^p} M_X(s)\right]_{s=0} = \left[\sum_{k=p}^{\infty} \frac{s^{k-p}}{p!}\,\mathbb{E}(X^k)\right]_{s=0} = \mathbb{E}(X^p) + \sum_{k=p+1}^{\infty} 0 = \mathbb{E}(X^p).$$

Consider a random variable Z with moment-generating function

$$M_Z(s) = \frac{a - 3s}{s^2 - 6s + 8} \qquad \text{for } |s| < 2.$$

Calculate the following quantities:

a. The numerical value of the parameter a.

b. $\mathbb{E}(Z)$.

c. $\mathrm{var}(Z)$.

Solution:

a. By definition, we know that $M_Z(s) = \mathbb{E}(e^{sZ})$, so we must have

$$M_Z(0) = \mathbb{E}(e^{0Z}) = 1 = \frac{a}{8},$$

from which it follows that $a = 8$.

b. We find the first moment as

$$\mathbb{E}(Z) = \left[\frac{d}{ds} M_Z(s)\right]_{s=0} = \left[\frac{2}{(4-s)^2} + \frac{1}{(2-s)^2}\right]_{s=0} = \frac{3}{8}.$$

c. We find that $\mathrm{var}(Z) = \frac{11}{64}$, where the second moment is

$$\mathbb{E}(Z^2) = \left[\frac{d^2}{ds^2} M_Z(s)\right]_{s=0} = \left[\frac{4}{(4-s)^3} + \frac{2}{(2-s)^3}\right]_{s=0} = \frac{5}{16}.$$

4. **Revisiting Proofs Using Transforms**

   a. Let $X \sim \text{Poisson}(\lambda)$ and $Y \sim \text{Poisson}(\mu)$ be independent. Calculate the MGF of $X+Y$, and use this to show that $X+Y \sim \text{Poisson}(\lambda+\mu)$.

   b. Calculate the MGF of $X \sim \text{Exponential}(\lambda)$, and use this to find all of the moments of X.

   c. Repeat the above part, but for $X \sim \mathcal{N}(0,1)$.

Solution:

a. The MGF of X is

$$\begin{aligned}\mathbb{E}(\exp(sX)) &= \sum_{x\in\mathbb{N}} \exp(sx)\frac{\lambda^x \exp(-\lambda)}{x!} \\ &= \exp(-\lambda)\sum_{x\in\mathbb{N}} \frac{(\lambda \exp s)^x}{x!} = \exp(\lambda(\exp s - 1)),\end{aligned}$$

which converges for all $s \in \mathbb{R}$. The MGF of $X+Y$ is

$$\begin{aligned}\mathbb{E}(\exp(s(X+Y))) = \mathbb{E}(\exp(sX)\cdot\exp(sY)) &= \mathbb{E}(\exp(sX))\cdot\mathbb{E}(\exp(sY)) \\ &= \exp(\lambda(\exp s - 1))\cdot\exp(\mu(\exp s - 1)) \\ &= \exp((\lambda+\mu)(\exp s - 1)),\end{aligned}$$

which we recognize as the MGF of a $\text{Poisson}(\lambda+\mu)$ random variable.

Remark: In general, it is not easy to argue that the MGF uniquely determines the probability distribution, which requires a few assumptions on the MGF itself, but we will not worry about these issues in this course.

b. We calculate

$$M_X(s) = \int_0^\infty \exp(sx)\lambda\exp(-\lambda x)\,dx = \lambda\int_0^\infty \exp(-(\lambda - s)x)\,dx = \frac{\lambda}{\lambda - s},$$

which converges for $s < \lambda$. Expanding M_X as a geometric series,

$$M_X(s) = \frac{1}{1-\frac{s}{\lambda}} = \sum_{k\in\mathbb{N}} \left(\frac{s}{\lambda}\right)^k,$$

as long as $|s| < \lambda$. Comparing the last expression with

$$\mathbb{E}(\exp(sX)) = \mathbb{E}\left(\sum_{k\in\mathbb{N}} \frac{(sX)^k}{k!}\right) = \sum_{k\in\mathbb{N}} \frac{s^k\,\mathbb{E}(X^k)}{k!}$$

and matching terms, we can argue that $\mathbb{E}(X^k) = \frac{k!}{\lambda^k}$.

c. The MGF of the standard Gaussian is

$$\begin{aligned}M_X(s) &= \int_{-\infty}^\infty \exp(sx)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)\,dx \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty \exp\left(-\frac{1}{2}(x^2 - 2sx)\right)\,dx\end{aligned}$$

$$= \exp\left(\frac{s^2}{2}\right)\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(-\frac{(x-s)^2}{2}\right)\,dx$$
$$= \exp\left(\frac{s^2}{2}\right).$$

Expanding as a power series,

$$M_X(s) = \sum_{k\in\mathbb{N}}\frac{s^{2k}}{2^k k!},$$

so by comparing terms as before, we see that $\mathbb{E}(X^k) = 0$ if k is odd, and

$$\mathbb{E}(X^k) = \frac{k!}{2^{k/2}(\frac{k}{2})!} = (k-1)!!$$

if k is even.

5. **Kelly strategy**

Consider a gambler who at each gamble either wins or loses his bet with probabilities p and $1-p$, independent of earlier gambles. When $p > 1/2$, a popular gambling system, known as the Kelly strategy, is to always bet the fraction $2p-1$ of the current fortune. Compute the expected fortune after n gambles, starting with x units and employing the Kelly strategy

Solution: If the gambler's fortune at the beginning of a round is a, the gambler bets $a(2p-1)$. He therefore gains $a(2p-1)$ with probability p, and loses $a(2p-1)$ with probability $1-p$. Thus, his expected fortune at the end of a round is

$$a(1+p(2p-1)-(1-p)(2p-1)) = a(1+(2p-1)^2).$$

Let X_k be the fortune after the kth round. Using the preceding calculation, we have

$$\mathbb{E}[X_{k+1}|X_k] = (1+(2p-1)^2)X_k.$$

Using the law of iterated expectations, we obtain

$$\mathbb{E}[X_{k+1}] = (1+(2p-1)^2)\mathbb{E}[X_k],$$

and

$$\mathbb{E}[X_1] = (1+(2p-1)^2)x.$$

We conclude that

$$\mathbb{E}[X_n] = (1+(2p-1)^2)^n x.$$

6. **Covariance Matrix**

For a random vector $X = [X_1, X_2, \ldots, X_n]^\mathsf{T}$, its covariance matrix Σ is defined with entries $\Sigma_{ij} = \mathrm{Cov}(X_i, X_j)$. Suppose that $\mathbb{E}[X] = 0$.

a. Show that Σ is positive semi-definite, i.e. for all $v \in \mathbb{R}^n$, we have $v^\mathsf{T}\Sigma v \geq 0$.

b. Show that if the X_i's are *pairwise* independent, then Σ is diagonal.

c. Give an example of two random variables X_1, X_2 with a diagonal covariance matrix, but such that X_1, X_2 are not independent.

Solution:

a. Note that we can write $\Sigma = \mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^\mathsf{T}]$. Therefore for any $v \in \mathbb{R}^n$, we have

$$v^\mathsf{T}\Sigma v = v^\mathsf{T}\,\mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^\mathsf{T}]v = \mathbb{E}[v^\mathsf{T} X X^\mathsf{T} v] = \mathbb{E}[(v^\mathsf{T} X)(v^\mathsf{T} X)^\mathsf{T}] \geq 0.$$

b. If X_i's are pairwise independent, then $\mathrm{Cov}(X_i, X_j) = 0$ for all $i \neq j$, and so Σ is diagonal.

c. Let $X_1 \sim \mathcal{N}(0, 1)$ and $X_2 = ZX_1$, where Z is uniform on $\{-1, 1\}$.