Problems: Gaussian random variables and a little bit of linear algebra

ISSYP 2016

1. *Mean and variance for biased coin flips.* Assume the following properties of random variables:

$$\begin{aligned}
\langle X_1 + X_2 \rangle &= \langle X_1 \rangle + \langle X_2 \rangle && \text{always} \\
\langle cX \rangle &= c\langle X \rangle && \text{if } c \text{ is constant (i.e. not a random variable)} \\
\langle X_1 X_2 \rangle &= \langle X_1 \rangle \langle X_2 \rangle && \text{if } X_1, X_2 \text{ are independent random variables}
\end{aligned} \tag{1}$$

(a) Using these properties, show that if $X = X_1 + X_2$, where X_1, X_2 are independent random variables, then $\mathrm{Var}(X) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2)$.

(b) Consider a biased coin which is heads with probability p, and tails with probability $(1-p)$. Let H_1 be the number of heads after a single flip (either 0 or 1). What are the mean and variance of the random variable H_1? Let H_N be the number of heads after N flips. What are the mean and variance of H_N?

2. *A toy example of the central limit theorem.* Define a random variable X by generating a random number θ between 0 and 2π, and then setting $X = \cos(\theta)$.

(a) Compute the mean $\bar{X}$ and variance $\mathrm{Var}(X)$ analytically. Hint: these can be written as integrals over θ.

(b) Now suppose we define a random variable $Z_N = (X_1 + \cdots + X_N)/N^{1/2}$. In the limit of large N, what probability distribution $p(Z_N)$ is predicted by the central limit theorem?

(c) Computer exercise: in a programming language of your choice, write a function which makes a random realization of the random variable X, and the random variable Z_{20}. By simulating many X's and taking an appropriate average, show that the mean and variance agree with what you calculated in part (a). By simulating many Z_{20}'s and making a histogram, show that the probability distribution agrees with what you calculated in part (b).

3. *Gaussian integrals.* (Warning: hard!) In this problem, you can assume that:

$$\int_{-\infty}^{\infty} dx\, e^{-x^2} = \sqrt{\pi} \tag{2}$$

(If you're curious how this is shown, there is a famous trick which is explained in the appendix!)

(a) By change of variable show that

$$\int_{-\infty}^{\infty} dx\, e^{-ax^2} = \sqrt{\pi} a^{-1/2} \tag{3}$$

where $a > 0$ is a real number.

(b) Now show that

$$\int_{-\infty}^{\infty} dx\, x^2 e^{-ax^2} = \frac{\sqrt{\pi}}{2} a^{-3/2} \tag{4}$$

using one of two possible approaches. One way is to group the integrand on the LHS as $(x)(xe^{-ax^2})$ and use integration by parts. The second way is to differentiate both sides of Eq. (3) with respect to a.

(c) Using the previous results show that the Gaussian probability distribution in one variable

$$p(x) = \frac{1}{(2\pi)^{1/2}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{5}$$

is correctly normalized (i.e. $\int p(x) = 1$) with variance σ^2 (i.e. $\int x^2 p(x) = \sigma^2$), as implicitly assumed in the lecture.

(d) Can you generalize part (b) to give a formula for $\int x^N e^{-ax^2}$, where N is a positive integer?

Appendix A: A famous trick for calculating $\int_{-\infty}^{\infty} dx e^{-x^2}$

As far as I know, the following strange trick is the only way of doing the integral! You'll need to have studied a little bit of multivariate calculus, in particular changing variables from Cartesian to polar coordinates in a 2D integral. Define:

$$I = \int_{-\infty}^{\infty} dx\, e^{-x^2} \tag{A1}$$

We can write I^2 as a 2D integral over the (x, y) plane.

$$\begin{aligned} I^2 &= \left(\int_{-\infty}^{\infty} dx\, e^{-x^2}\right)\left(\int_{-\infty}^{\infty} dy\, e^{-y^2}\right) \\ &= \int\int dx\, dy\, e^{-x^2-y^2} \end{aligned} \tag{A2}$$

This doesn't appear to be making progress, but if we now change variables to polar coordinates (r, θ), then we can do the integral! (In the steps below, we integrate over θ first and then r.)

$$\begin{aligned} I^2 &= \int dr\, d\theta\, re^{-r^2} \\ &= 2\pi \int_0^{\infty} dr\, re^{-r^2} \\ &= 2\pi \left(\frac{1}{2}\right) \end{aligned} \tag{A3}$$

Note that the extra factor of r we picked up in the change of variables to polar coordinates is what allows us to do the integral. Taking the square root on both sides we now get $I = \sqrt{\pi}$ as desired.