Probability and Combinatorics Worksheet 6

Questions

These are the questions you should make sure you work on in the workshop.

1. The distribution of tree heights in a christmas tree forest is

$$p(h) = \begin{cases} 0.3 & 0 \leq h < 2 \\ 0.2 & 2 \leq h < 4 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

What is the mean height of trees in the forest?

Solution: So

$$\mu = \langle H \rangle = \int_{-\infty}^{\infty} hf(h) = \int_0^2 0.3hdh + \int_2^4 0.2hdh \quad (2)$$

and

$$\int_0^2 0.3hdh = 0.3\frac{h^2}{2}\bigg|_0^2 = 0.6 \quad (3)$$

and

$$\int_2^4 0.2hdh = 0.2\frac{h^2}{2}\bigg|_2^4 = 0.2(8-2) = 1.2 \quad (4)$$

so $\mu = 1.8$.

2. Work out the mean and variance for the distribution

$$p(x) = \begin{cases} 2/a & x \in [-a, a] \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

Solution: So in this case

$$\langle X \rangle = \int_{-\infty}^{\infty} xp(x)dx = \frac{2}{a}\int_{-a}^{a} xdx = 0 \quad (6)$$

and

$$\langle X^2 \rangle = \int_{-\infty}^{\infty} x^2p(x)dx = \frac{2}{a}\int_{-a}^{a} x^2dx = \frac{2x^3}{3a}\bigg]_{-a}^{a} = \frac{4a^2}{3} \quad (7)$$

and here the variance is the same as the second moment because the mean is zero.

3. Starting from the expression for the mean

$$\mu = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} xe^{-\frac{(x-\mu)^2}{2\sigma^2}}dx \quad (8)$$

show the Gaußian distribution has variance σ^2.

Solution: Differenciate with respect to μ to get

$$1 = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x \frac{x-\mu}{\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \tag{9}$$

and multiply across by the σ^2 to get

$$\sigma^2 = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x(x-\mu) e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \tag{10}$$

and then a bit of algebra gives

$$\sigma^2 = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx - \mu \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \tag{11}$$

4. The size of a standard croquet ball is 3 5/8 inches[1]. The height of a croquet hoop is 3 3/4 inches. If a not very good croquet-ball making machine makes croquet balls whose mean matches the standard and with standard deviation 1/8 inch, what is the chance it will make a ball too large to fit through the hoop? You can write the solution in terms of the error function.

Solution: So

$$z = \frac{x-\mu}{\sqrt{2}\sigma} \tag{12}$$

so for $x_1 = 3.75$ in, we have

$$z_1 = \frac{1/8}{\sqrt{2}/8} = \frac{1}{\sqrt{2}} \tag{13}$$

Any height bigger than this will not fit, so $z_2 = \infty$ and $\text{erf}\infty = 1$ so

$$\text{Prob}(x > 3.75) = \frac{1}{2}[1 - \text{erf}(1/\sqrt{2})] \approx 0.16 \tag{14}$$

where the 0.16 is given for interest, it wasn't expected as part of the answer.

Extra questions

1. This will look like a long question but it is almost all background and the question is not too bad when you actually read through it. In particle physics when a collider is being used to find a new particle like the Higgs boson or the top squark scientists don't detect the sought after particle directly since it usually decays almost straight away, instead they detect the more common particles that particle will decay into, for example, a Higgs boson can decay in to two photons and these can be detected. Roughly speaking scientists count these events. However, the whole situation is very messy and there will always be some events even if the particle doesn't exist at the energy being examined. The amount of these background events will fluctuate from experiment to experiment, typically like a Gaußian. The scientific team is allowed to claim they have discovered

[1] Everything in croquet is measured in old timey units

the particle if the number of events they measure is more than five standard deviations above what would be expected if the particle didn't exist. What is the probability of this 'discovery' happening by chance?

Solution: So we are interested in the probability of a results bigger than $\mu + 5\sigma$. Now

$$z_1 = \frac{\mu + 5\sigma - \mu}{\sqrt{2}\sigma} = \frac{5}{\sqrt{2}} \tag{15}$$

and

$$\text{Prob}(x > \mu + 5\sigma) = \frac{1}{2}[1 - \text{erf}(5/\sqrt{(2)})] \tag{16}$$

which is about one chance in 3.5 million.

2. Another useful distribution is the exponential distribution:

$$p(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

What is the probability $\text{Prob}(x_1 < x < x_2)$ where x_1 and x_2 are both positive.

Solution:

$$\text{Prob}(x_1 < x < x_2) = \lambda \int_{x_1}^{x_2} e^{-\lambda y} dy = -\left. e^{-\lambda y} \right]_{x_1}^{x_2} = e^{-\lambda x_1} - e^{-\lambda x_2} \tag{17}$$

3. What is the mean of the exponential distribution?

Solution: Well

$$1 = \lambda \int_0^\infty e^{-\lambda x} dx \tag{18}$$

and differentiate both sides with respect to λ to get

$$0 = \int_0^\infty e^{-\lambda x} dx - \lambda \int_0^\infty x e^{-\lambda x} dx \tag{19}$$

so

$$\langle X \rangle = \frac{1}{\lambda} \tag{20}$$

4. Australorp hens weigh on average 4kg with a standard deviation 0.25kg; in one farm australorps who weigh less than 3.5kg are fed *patent chicken spicer*, a mixture of chalk, corn and pepper. What fraction of these hens are fed patent chicken spicer?

Solution: Here is a picture of one of my hens, an Australorp, being fed a raspberry.

So we are asking for $P(x < 3.5)$ and since

$$z = \frac{3.5 - 4}{0.25\sqrt{2}} = \sqrt{2}$$

the probability, and hence the proportion, is

$$P(x < 3.5) = \frac{1}{2}[1 - \text{erf}\,(\sqrt{2})] \approx 0.02275$$