MTH5129 Probability & Statistics II

Coursework 1

1. In this question we revise some standard probability distributions (i.e. random variables). For each of the following distributions write down its *probability mass function*, its *mean*, its *variance* and a *description of the experiment* related to it.

   a) **Bernoulli distribution** with parameter p

   b) **Binomial distribution** with parameters n and p

   c) **Geometric distribution** with parameter p

   d) **Poisson distribution** with parameter λ

   Solution: This is standard material from any textbook on probability. However, we present below a broad description of each distribution to avoid confusion:

   a) **Bernoulli distribution** with parameter p – i.e. a single trial with probability p of success.

   b) **Binomial distribution** with parameters N and p – i.e. N independent trials each of which has probability p of success.

   c) **Geometric distribution** with parameter p – i.e. a sequence of independent trials each with probability of success p, where the random variable counts the number of trials until the first success. *This definition includes the first successful trial; others may sometimes just count the number of failures.*

   d) **Poisson distribution** with parameter λ — a sequence of independent events each occurring at a rate λ, where the random variable counts the number of events that occurred in a particular time interval.

2. In this question, we move on to the revision of some "common" continuous random variables

   a) **Uniform Distribution.** A random variable X has Uniform distribution on $[a, b]$ and write $X \sim U(a, b)$ if

$$f_X(x) = \begin{cases} \frac{1}{(b-a)} & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases}$$

   Prove that this probability density function integrates to one.

Solution: Note that all intervals of equal length within the support $[a, b]$ of the probability density function have equal probability of occurrence: if $a \leq \alpha < \beta \leq b$ then

$$P(X \in (\alpha, \beta)) = \int_{\alpha}^{\beta} \frac{1}{b-a}\, dx = \frac{\beta - \alpha}{b - a}.$$

The verification of the probability density function follows.

b) **Exponential Distribution.** A random variable X has exponential distribution with parameter $\lambda > 0$ and write $X \sim Exp(\lambda)$ if

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

Prove that this probability density function integrates to one.

Solution: We know f_X, but it is defined with multiple "cases". Therefore we use an important "trick" to deal with this. We know that

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

(this would be true whatever f_X was).

In the second integral x ranges from 0 to ∞: for all of these values of x we know that $f_X(x) = e^{-x}$ (we are always in the first case of the definition of f_x) so we can replace $f_X(x)$ by e^{-x} in the second integral. Similarly, in the first integral x ranges from $-\infty$ to 0 so $f_X(x) = 0$ (we are always in the second case of the definition of f_X) so we can replace $f_X(x)$ by 0 in the first integral.

Putting this together we get

$$\begin{aligned} \int_{-\infty}^{\infty} f_X(x)\, dx &= \int_{-\infty}^{0} f_X(x)\, dx + \int_{0}^{\infty} f_X(x)\, dx \\ &= \int_{-\infty}^{0} 0\, dx + \int_{0}^{\infty} e^{-x}\, dx \\ &= 0 + \left[-e^{-x}\right]_{x=0}^{\infty} = e^{0} = 1 \end{aligned}$$

as required.

Remark 1 (IMPORTANT)

– *If you are given a probability density function, it means that the random variable is continuous, so you can integrate to find probabilities.*

– *The above example is very important. Almost all probability density functions have multiple cases in their definition and you need to be able to integrate them: the whole point of the probability density function is that you can integrate to get a probability.*

c) **Gamma Distribution.** A random variable X has a Gamma distribution with *shape* parameter $\alpha > 0$ and *rate* parameter $\beta > 0$ and we write $X \sim Ga(\alpha, \beta)$ if

$$f_X(x) = \begin{cases} \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} & \text{if } x > 0, \\ 0 & \text{if } x \leq 0, \end{cases}$$

where $\Gamma(\alpha)$ is the Gamma function, which is given by

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \, dx.$$

Prove that this probability density function integrates to one.

Solution: We use a simple change of variable $y = \beta x$ in the integral and the verification of the probability density function follows.

Remark 2 *Note that in some textbooks, the Gamma distribution can be defined with $\beta = 1/\theta$ for the scale parameter θ instead, or write it as $Ga(\beta, \alpha)$. You always have to make sure what each parameter stands for and know what notation is used in each case.*

d) **Chi-Square.** A random variable X has a Chi-Square distribution with ν *degrees of freedom* and we write $X \sim \chi^2(\nu)$, if X has a $Ga(\nu/2, 1/2)$ distribution for some integer $\nu \in \mathbb{N}$. In such a case,

$$f_X(x) = \begin{cases} \frac{x^{\nu/2-1} \, e^{-x/2}}{2^{\nu/2} \, \Gamma(\nu/2)} & \text{if } x > 0, \\ 0 & \text{if } x \leq 0, \end{cases}$$

Using the above probability density function, and the fact that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ (not shown here), it is easy to see that, the probability density function of $X \sim \chi^2(1)$ is

$$f_X(x) = \begin{cases} \frac{1}{\sqrt{2\pi x}} \, e^{-x/2} & \text{if } x > 0, \\ 0 & \text{if } x \leq 0, \end{cases}$$

Prove that this probability density function integrates to one.

Solution: Proving that these probability density functions integrate to one should follow similar arguments to the ones for the *Gamma* distribution, but you may try it for exercise.

e) **Normal Distribution.** A random variable X has Normal distribution with parameters (μ, σ), for $\sigma > 0$, and write $X \sim N(\mu, \sigma^2)$ if

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

We shall use the fact that the Normal probability density function integrates to 1 (the proof is a bit technical – involves passing to polar coordinates – not shown in this course – non-examinable). It is however useful to observe that by a simple change of variable in the integral, namely $y = (x - \mu)/\sigma$ we obtain

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

We now see that the integral in the left hand side of the above formula does not depend on the parameters (μ, σ)!

f) **(Student's) t Distribution.** A random variable X has t distribution with ν degrees of freedom and we write $X \sim t_\nu$ if

$$f_X(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}.$$

We shall use the fact that the above probability density function integrates to 1 (non-examinable).

Remark 3 *Some books call this distribution Student's t, because the first person to derive it published his result under the pseudonym of "Student".*

g) **Cauchy Distribution.** A random variable X has Cauchy distribution with *location* parameter x_0 (specifying the location of the peak of the distribution) and *scale* parameter γ and we write $X \sim Cauchy(x_0, \gamma)$ if

$$f_X(x) = \frac{1}{\pi\gamma} \left[\frac{\gamma^2}{(x - x_0)^2 + \gamma^2}\right],$$

It is an unusual distribution because it has no defined mean or variance (we will see why in the lecture notes).

Moreover if X is Cauchy then $1/X$ is also Cauchy (we will learn several methods for proving such type of statements).

3. The Gamma function Γ involved in $Ga(\alpha, \beta)$ distribution is given by

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x}\, dx.$$

(i). Prove that, for any $\alpha > 0$, we have

$$\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1).$$

(ii). Prove that, for any integer $n \geq 1$, we have

$$\Gamma(n) = (n - 1)!.$$

Solution:

(i). We have from the definition of $\Gamma(\alpha)$ that

$$\Gamma(\alpha-1)=\int_0^\infty x^{\alpha-2}e^{-x}\,dx$$

Using the fact that

$$\frac{d}{dx}\left(\frac{1}{\alpha-1}x^{\alpha-1}\right)=x^{\alpha-2},$$

we get from Integration by Parts, that

$$\begin{aligned}
\Gamma(\alpha-1) &= \int_0^\infty x^{\alpha-2}e^{-x}\,dx \\
&= \int_0^\infty \frac{d}{dx}\left(\frac{1}{\alpha-1}x^{\alpha-1}\right)e^{-x}\,dx \\
&= \left[\frac{1}{\alpha-1}x^{\alpha-1}e^{-x}\right]_0^\infty - \int_0^\infty \frac{1}{\alpha-1}x^{\alpha-1}\frac{d}{dx}\left(e^{-x}\right)\,dx \\
&= 0-0+\frac{1}{\alpha-1}\int_0^\infty x^{\alpha-1}e^{-x}\,dx \\
&= \frac{1}{\alpha-1}\Gamma(\alpha)
\end{aligned}$$

which implies the desired formula.

(ii). Use the definition to calculate that $\Gamma(1)=1$. Then use the above part, to prove (e.g. by induction – see your Number, Sets & Functions notes) the desired formula.

4. Given a normal random variable $X\sim N(\mu,\sigma^2)$, prove that $E(X)=\mu$ and $\mathrm{Var}(X)=\sigma^2$.

Solution: By the definition of a normal random variable the probability density function of X is

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

By the definition of expectation we have:

$$\begin{aligned}
E(X) &= \int_{-\infty}^\infty xf_X(x)dx=\int_{-\infty}^\infty x\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx \\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^\infty(\sigma u+\mu)e^{-\frac{u^2}{2}}\sigma du.
\end{aligned}$$

where the last step is due to the following change of variable: $u=\frac{x-\mu}{\sigma}$. Then $x=\sigma u+\mu$, $dx=\sigma du$, and the region of integration remains unchanged. Thus

$$E(X)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty(\sigma u+\mu)e^{-\frac{u^2}{2}}du=\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^\infty ue^{-\frac{u^2}{2}}du+\mu\frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-\frac{u^2}{2}}du.$$

Finally, $\int_{-\infty}^{\infty} ue^{-\frac{u^2}{2}}du = 0$ because the function $g(u) = ue^{-\frac{u^2}{2}}$ is odd and $\int_{-\infty}^{\infty} e^{-\frac{u^2}{2}}du = \sqrt{2\pi}$ (see the definition of the normal random variable and the computation explained there).

We thus obtain that $E(X) = \mu$.

To compute the variance, we use the definition of variance and the above Theorem to get

$$\mathrm{Var(X) = E(X - E(X))^2 = E(X - \mu)^2} = \int_{-\infty}^{\infty} \mathrm{(x - \mu)^2 f_X(x)dx.}$$

The same change of variable gives

$$\begin{aligned}\int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (\sigma u)^2 e^{-\frac{u^2}{2}} \sigma du \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u^2 e^{-\frac{u^2}{2}} du.\end{aligned}$$

To compute the last integral we use the integration by parts:

$$\int_{-\infty}^{\infty} u^2 e^{-\frac{u^2}{2}} du = -\int_{-\infty}^{\infty} u\, d(e^{-\frac{u^2}{2}}) = -ue^{-\frac{u^2}{2}}|_{u=-\infty}^{u=\infty} + \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du = \sqrt{2\pi}.$$

We thus obtain that $\mathrm{Var}(X) = \sigma^2$.

Remark 4 *In the above calculation, we use* $ue^{-\frac{u^2}{2}}|_{u=-\infty}^{u=\infty} = 0$. *The proof of this simple fact goes as follows:*

$$\begin{aligned}ue^{-\frac{u^2}{2}}|_{u=-\infty}^{u=\infty} &= \lim_{M\to-\infty,\ N\to\infty} ue^{-\frac{u^2}{2}}|_{u=M}^{u=N} = \lim_{M\to-\infty,\ N\to\infty} (Ne^{-\frac{N^2}{2}} - Me^{-\frac{M^2}{2}}) \\ &= \lim_{N\to\infty} Ne^{-\frac{N^2}{2}} - \lim_{M\to-\infty} Me^{-\frac{M^2}{2}} = 0,\end{aligned}$$

where the latter follows from the L'Hôpital's rule, e.g.

$$\lim_{N\to\infty} Ne^{-\frac{N^2}{2}} = \lim_{N\to\infty} \frac{N}{e^{\frac{N^2}{2}}} = \lim_{N\to\infty} \frac{1}{Ne^{\frac{N^2}{2}}} = 0.$$

5. Suppose that there are two urns. The first contains 5 red balls, 3 green balls, and 2 blue balls. The second contains 2 red balls and 4 green balls.

   We pick a ball at random from the first urn and place it in the second urn. We then pick a ball at random from the second urn (which might be the ball we have just placed there).

   a) What is the probability this ball is red?

   b) What is the probability it is green?

c) What is the probability it is blue?

d) What is the expected number of trials of the above experiment until we finally pick a blue ball?

Solution:

a) Let A be the event that the ball we pick at the end is red, let B_1 be the event that the ball we pick from the first urn is red, let B_2 be the event that the first ball we pick from the first urn is green and let B_3 be the event that the first ball we pick from the first urn is blue. Obviously B_1, B_2 and B_3 form a partition. Thus we use the Theorem of Total Probability.

We see that $P(B_1) = 5/10 = 1/2$, that $P(B_2) = 3/10 = 1/5$ and that $P(B_3) = 2/10$.

Now if B_1 occurs then there are 3 red balls and 4 green balls in the second urn when we pick, so $P(A \mid B_1) = 3/7$.

If B_2 occurs then there are 2 red balls and 5 green balls in the second urn so $P(A \mid B_2) = 2/7$.

If B_3 occurs then there are 2 red balls 4 green balls and a blue ball in the second urn so $P(A \mid B_3) = 2/7$.

Thus by the Theorem of Total Probability

$$\begin{aligned} P(A) &= P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + P(A \mid B_3)P(B_3) \\ &= \frac{5}{10} \cdot \frac{3}{7} + \frac{3}{10} \cdot \frac{2}{7} + \frac{2}{10} \cdot \frac{2}{7} = \frac{25}{70} = \frac{5}{14} \end{aligned}$$

b) For the second part, let A be the event that the ball we pick at the end is green. We use the same partition as in the first part.

As above we calculate the conditional probabilities. This time $P(A \mid B_1) = 4/7$, $P(A \mid B_2) = 5/7$ and $P(A \mid B_3) = 4/7$. Thus by the Theorem of Total Probability

$$\begin{aligned} P(A) &= P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + P(A \mid B_3)P(B_3) \\ &= \frac{5}{10} \cdot \frac{4}{7} + \frac{3}{10} \cdot \frac{5}{7} + \frac{2}{10} \cdot \frac{4}{7} = \frac{43}{70}. \end{aligned}$$

c) For the final part, let A be the event that the ball we pick at the end is blue. We use the same partition as in the first part.

As above we calculate the conditional probabilities. This time $P(A \mid B_1) = 0$, $P(A \mid B_2) = 0$ and $P(A \mid B_3) = 1/7$. Thus by the Theorem of Total Probability

$$\begin{aligned} P(A) &= P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + P(A \mid B_3)P(B_3) \\ &= \frac{5}{10} \cdot 0 + \frac{3}{10} \cdot 0 + \frac{2}{10} \cdot \frac{1}{7} = \frac{2}{70} = \frac{1}{35}. \end{aligned}$$

d) Let X be a random variable counting the number of experiments performed until we finally pick a blue ball. We call an experiment to be a "success" if we choose a blue ball at the end. This is a Geometric random variable with probability of success given by the part above, namely $p = 1/35$.

Therefore, $X \sim Geom(p)$ and we know that $E(X) = 1/p = 35$ (Prove this – Check your Introduction to Probability notes). This means that we expect to perform the experiment 35 times in order to finally pick a blue ball.

6. I roll two fair dice. Use the Theorem of Total Probability for Expectations to calculate the expected value of the product of the two numbers rolled.

Solution: Let X be the product of the two dice. Let $B_1, B_2, \ldots, B_6$ be the events that the first die shows a $1, 2, \ldots, 6$ respectively. These obviously form a partition and $P(B_i) = 1/6$ for each i.

Now define by D the outcome of the second die and calculate

$$E(X \mid B_i) = E(i\,D) = i\,E(D) = \frac{7}{2}i.$$

Hence,

$$E(X) = \sum_{i=1}^{6} \frac{7}{2}i \cdot \frac{1}{6} = \frac{7}{12}\sum_{i=1}^{6} i = \frac{7}{12}\frac{6 \cdot 7}{2} = \frac{49}{4}.$$