Problem Set: Continuous Random Variables

James Heald
Gatsby Bridging Programme 2023

Problem 1. Suppose X is a continuous random variable whose probability density function is given by

$$f_X(x) = \begin{cases} C(4x - 2x^2) & \text{if } 0 < x < 2 \\ 0 & \text{otherwise.} \end{cases}$$

(a) What is the value of C?

(b) Find $P(X > 1)$.

Problem 2. The amount of time in hours that a computer functions before breaking down is a continuous random variable with probability density function given by

$$f_X(x) = \begin{cases} \lambda e^{-x/100} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases}$$

What is the probability that

(a) a computer will function between 50 and 150 hours before breaking down?

(b) it will function for fewer than 100 hours?

Problem 3. Buses arrive at a specified stop at 15-minute intervals starting at 7 A.M. That is, they arrive at 7, 7:15, 7:30, 7:45, and so on. If a passenger arrives at the stop at a time that is uniformly distributed between 7 and 7:30, find the probability that they wait

(a) less than 5 minutes for a bus;

(b) more than 10 minutes for a bus.

Problem 4. In commuting to work, a professor must first get on a bus near her house and then transfer to a second bus. If the waiting time (in minutes) at each stop has a uniform distribution on the interval $[a, b]$, then it can be shown that the total waiting time X has the probability density function

$$f_X(x) = \begin{cases} \frac{x}{25} & \text{if } 0 \leq x < 5 \\ \frac{2}{5} - \frac{x}{25} & \text{if } 5 \leq x \leq 10 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Obtain an expression for the $(100p)$th percentile $\eta(p)$ [Hint: consider $0 \leq \eta(p) < 5$ and $5 \leq \eta(p) \leq 10$ separately].

(b) Find the median.

Problem 5. The probability density function of a continuous random variable X is given by

$$f_X(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find $\mathbb{E}[e^X]$.

Problem 6. Suppose the probability density function of the magnitude X of a dynamic load on a bridge (in Newtons) is

$$f_X(x) = \begin{cases} \frac{3x}{8} + \frac{1}{8} & \text{if } 0 \le x \le 2 \\ 0 & \text{otherwise.} \end{cases}$$

Find

(a) $\mathbb{E}[X]$;

(b) $\mathrm{Var}[X]$.

Problem 7. The breakdown voltage of a randomly chosen diode is a random variable X that is known to be normally distributed. Using Table 1, find the probability that a diode's breakdown voltage is within 1.5 standard deviations of its mean value.

AREA $\Phi(x)$ UNDER THE STANDARD NORMAL CURVE TO THE LEFT OF X

X	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Table 1

Problem 8. X is a normal random variable with parameters $\mu = 3$ and $\sigma^2 = 9$. Using Table 1, find

(a) $P(2 < X < 5)$;

(b) $P(X > 0)$;

(c) $P(|X - 3| > 6)$.

Problem 9. The continuous random variable X is said to have a Weibull distribution with scale parameter $k > 0$ and scale parameter $\sigma > 0$ if the cumulative distribution function of X is

$$F_X(x) = \begin{cases} 1 - e^{-(x/\sigma)^k} & \text{if } x \ge 0 \\ 0 & \text{otherwise.} \end{cases}$$

Find a function Φ such that the random variable $Y = \Phi(U)$ has the same distribution as X, where U is a continuous random variable having a uniform distribution on the interval $[0, 1]$.

Problem 10. The continuous random variable X is said to have a pareto distribution with scale parameter $\beta > 0$ and shape parameter $\alpha > 0$ if the probability density function of X is

$$f_X(x) = \begin{cases} \frac{\alpha\beta^\alpha}{x^{\alpha+1}} & \text{if } \beta \leq x < \infty \\ 0 & \text{otherwise.} \end{cases}$$

Find a function Φ such that the random variable $Y = \Phi(U)$ has the same distribution as X, where U is a continuous random variable having a uniform distribution on the interval $[0, 1]$.

Solution: Problem 1.

(a) Since f_X is a probability density function, we must have $\int_{-\infty}^{\infty} f_X(x)dx = 1$, implying that

$$C\int_0^2 (4x - 2x^2)dx = 1$$
$$C\left[2x^2 - \frac{2x^3}{3}\right]\Bigg|_{x=0}^{x=2} = 1,$$

and hence $C = \frac{3}{8}$.

(b) $P(X > 1) = \int_1^{\infty} f_X(x)dx = \frac{3}{8}\int_1^2 (4x - 2x^2)dx = \frac{1}{2}$.

Solution: Problem 2.

(a) First find the value of λ such that $\int_{-\infty}^{\infty} f_X(x)dx = 1$.

$$\lambda\int_0^{\infty} e^{-x/100}dx = 1$$
$$-\lambda(100)e^{-x/100}\Bigg|_{x=0}^{x=\infty} = 1$$
$$\lambda(100) = 1,$$

which implies that $\lambda = 1/100$. Hence,

$$\begin{aligned} P(50 < X < 150) &= \int_{50}^{150} \frac{1}{100}e^{-x/100}dx \\ &= -e^{-x/100}\Bigg|_{50}^{150} \\ &= e^{-1/2} - e^{-3/2} \\ &\approx 0.384 \end{aligned}$$

(b) Similarly, $P(X < 100) = \int_0^{100} \frac{1}{100}e^{-x/100}dx = -e^{-x/100}\Big|_0^{100} = 1 - e^{-1} \approx 0.633$. In other words, approximately 63.3 percent of the time, a computer will fail before registering 100 hours of use.

Solution: Problem 3.

Let X denote the number of minutes past 7 that the passenger arrives at the stop. Since X is a uniform random variable over the interval (0, 30), it follows that the passenger will have to wait less than 5 minutes if (and only if) they arrive between 7:10 and 7:15 or between 7:25 and 7:30. Hence, the desired probability for part (a) is

$$P(10 < X < 15) + P(25 < X < 30) = \int_{10}^{15} \frac{1}{30}dx + \int_{25}^{30} \frac{1}{30}dx = 1/3.$$

Similarly, they would have to wait more than 10 minutes if they arrive between 7 and 7:05 or between 7:15 and 7:20, so the probability for part (b) is

$$P(0 < X < 5) + P(15 < X < 20) = \int_0^5 \frac{1}{30}dx + \int_{15}^{20} \frac{1}{30}dx = 1/3.$$

Why are the answers to (a) and (b) the same?

Solution: Problem 4.

(a) The $(100p)$th percentile is the value of $\eta(p)$ that satisfies $p = \int_{-\infty}^{\eta(p)} f_X(x)dx$. It is helpful to first calculate the probability that the total waiting time is less than 5 minutes:

$$\begin{aligned}
P(0 \leq X < 5) &= \int_0^5 \frac{x}{25} dx \\
&= \frac{1}{50} x^2 \Big|_0^5 \\
&= 0.5.
\end{aligned}$$

Therefore, $0 \leq \eta(p) < 5$ when $0 \leq p \leq 0.5$. In this scenario, the $(100p)$th percentile is the value of $\eta(p)$ that satisfies

$$\begin{aligned}
p &= \int_0^{\eta(p)} \frac{x}{25} dx \\
&= \frac{1}{50} x^2 \Big|_0^{\eta(p)},
\end{aligned}$$

which is $\eta(p) = \sqrt{50p}$. In contrast, $5 \leq \eta(p) \leq 10$ when $0.5 < p \leq 1$. In this scenario, the $(100p)$th percentile is the value of $\eta(p)$ that satisfies

$$\begin{aligned}
p &= 0.5 + \int_5^{\eta(p)} \left(\frac{2}{5} - \frac{x}{25}\right) dx \\
&= 0.5 + \left[\frac{2x}{5} - \frac{x^2}{50}\right]\Bigg|_5^{\eta(p)} \\
&= 0.5 + \left(\frac{2\eta(p)}{5} - \frac{\eta(p)^2}{50}\right) - \left(2 - \frac{1}{2}\right).
\end{aligned}$$

Here, the term 0.5 has been added to the right-hand side to account for the probability that the total waiting time is less than 5 minutes. After some algebraic rearranegment, the quadratic formula can be used to solve the equation $\eta(p)^2 - 20\eta(p) + 50p + 50 = 0$. This gives $\eta(p) = 10 - \sqrt{50(1-p)}$. Thus, the expression for the $(100p)$th percentile is

$$\eta(p) = \begin{cases} \sqrt{50p} & \text{if } 0 \leq p \leq 0.5 \\ 10 - \sqrt{50(1-p)} & \text{if } 0.5 < p \leq 1. \end{cases}$$

(b) The median is $\eta(0.5) = 5$.

Solution: Problem 5.

The expected value of e^X is

$$\begin{aligned}
\mathbb{E}[e^X] &= \int_{-\infty}^{\infty} e^X f_X(x) dx \\
&= \int_0^1 e^X dx \\
&= e^X \Big|_0^1 \\
&= e - 1.
\end{aligned}$$

Solution: Problem 6.

(a) The expected value of X is

$$\begin{aligned}\mathbb{E}[X] &= \int_0^2 x\left(\frac{3x}{8}+\frac{1}{8}\right)dx \\ &= \left[\frac{x^3}{8}+\frac{x^2}{16}\right]\Bigg|_0^2 \\ &= \frac{5}{4}.\end{aligned}$$

(b) The variance of X is

$$\begin{aligned}\text{Var}[X] &= \mathbb{E}[X^2]-\mathbb{E}[X]^2 \\ &= \int_0^2 x^2\left(\frac{3x}{8}+\frac{1}{8}\right)dx-\left(\frac{5}{4}\right)^2 \\ &= \left[\frac{3x^4}{32}+\frac{x^3}{24}\right]\Bigg|_0^2-\left(\frac{5}{4}\right)^2 \\ &= \frac{13}{48}.\end{aligned}$$

Solution: Problem 7.

This question can be answered without knowing either μ or σ, as long as the distribution is known to be normal; the answer is the same for any normal distribution:

$$\begin{aligned}P(X \text{ is within 1.5 standard deviations from the mean}) &= P(\mu-1.5\sigma \le X \le \mu+1.5\sigma) \\ &= P\left(\frac{\mu-1.5\sigma-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{\mu+1.5\sigma-\mu}{\sigma}\right) \\ &= P(-1.5 \le Z \le 1.5) \\ &= \Phi(1.5)-\Phi(-1.5) \\ &= \Phi(1.5)-[1-\Phi(1.5)] \\ &\approx 0.8664,\end{aligned}$$

where $Z=(X-\mu)/\sigma$ is a standard normal random variable with cumulative distribution function Φ.

Solution: Problem 8.

(a)

$$\begin{aligned}P(2<X<5) &= P\left(\frac{2-\mu}{\sigma}<\frac{X-\mu}{\sigma}<\frac{5-\mu}{\sigma}\right) \\ &= P\left(\frac{2-3}{3}<\frac{X-3}{3}<\frac{5-3}{3}\right) \\ &= P\bigg(-0.33<Z<0.66\bigg) \\ &= \Phi(0.66)-\Phi(-0.33) \\ &= \Phi(0.66)-[1-\Phi(0.33)] \\ &\approx 0.3747.\end{aligned}$$

(b)

$$\begin{aligned}
P(X > 0) &= P\left(\frac{X-3}{3} > \frac{0-3}{3}\right) \\
&= P\left(Z > -1\right) \\
&= 1 - P\left(Z \leq -1\right) \\
&= 1 - \Phi(-1) \\
&= 1 - [1 - \Phi(1)] \\
&\approx 0.8413.
\end{aligned}$$

(c)

$$\begin{aligned}
P(|X-3| > 6) &= P(X-3 > 6) + P(X-3 < -6) \\
&= P(X > 9) + P(X < -3) \\
&= P\left(\frac{X-3}{3} > \frac{9-3}{3}\right) + P(\frac{X-3}{3} < \frac{-3-3}{3}\right) \\
&= P(Z > 2) + P(Z < -2) \\
&= 1 - P(Z \leq 2) + P(Z < -2) \\
&= 1 - \Phi(2) + [1 - \Phi(2)] \\
&\approx 0.0456.
\end{aligned}$$

Solution: Problem 9.

The probability integral transform tells us that $\Phi = F_X^{-1}$, the inverse of the cumulative distribution function. Therefore, to find $\Phi = F_X^{-1}$, set $U = F_X(x)$ and solve for x as a function of U. This gives $\Phi = \sigma(-\log(1 - U))^{1/k}$.

Solution: Problem 10.

First find the cumulative distribution function F_X and then find its inverse F_X^{-1}.

$$\begin{aligned}
F_X(x) &= \int_{-\infty}^{x} f_X(t)dt \\
&= \int_{-\infty}^{x} \frac{\alpha\beta^\alpha}{t^{\alpha+1}} dt \\
&= \frac{-\beta^\alpha}{t^\alpha}\bigg|_\beta^x \\
&= 1 - \left(\frac{\beta}{x}\right)^\alpha.
\end{aligned}$$

To find $\Phi = F_X^{-1}$, set $U = F_X(x)$ and solve for x as a function of U. This gives $\Phi = \beta(1-U)^{-1/\alpha}$.