Probability and Statistics: MA6.101
Tutorial 4

Topics Covered: Continuous Random Variable, Functions of Random Variable

Q1: A continuous random variable X is given distribution with parameters $n, \gamma > 0$ with PDF given as:

$$f_X(x) = \begin{cases} cx^{n-1}e^{-\lambda x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find c.

(b) Find $\mathbb{E}[X]$ and $\text{Var}[X]$.

(c) If I write the values calculated above as a sum of n i.i.d. random variables, then what **could be a potential continuous random variable** that can sum to this random variable?

Hint: Simplify in terms of the gamma function:

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

A:

(a) We require $\displaystyle\int_0^\infty f_X(x)\,dx = 1$:

$$1 = \int_0^\infty c\,x^{n-1}e^{-\lambda x}\,dx = c\int_0^\infty x^{n-1}e^{-\lambda x}\,dx.$$

Use the change of variable $t = \lambda x$ so $x = t/\lambda$ and $dx = dt/\lambda$:

$$\int_0^\infty x^{n-1}e^{-\lambda x}\,dx = \int_0^\infty \left(\frac{t}{\lambda}\right)^{n-1} e^{-t}\,\frac{dt}{\lambda} = \frac{1}{\lambda^n}\int_0^\infty t^{n-1}e^{-t}\,dt = \frac{\Gamma(n)}{\lambda^n}.$$

Hence

$$1 = c\,\frac{\Gamma(n)}{\lambda^n} \quad\Longrightarrow\quad \boxed{c = \frac{\lambda^n}{\Gamma(n)}}$$

(For integer n, $\Gamma(n) = (n-1)!$, so $c = \lambda^n/(n-1)!$.)

(b) With $c = \lambda^n/\Gamma(n)$,

$$\mathbb{E}[X] = \int_0^\infty x\,f_X(x)\,dx = \frac{\lambda^n}{\Gamma(n)}\int_0^\infty x^n e^{-\lambda x}\,dx$$

Again set $t = \lambda x$:

$$\int_0^\infty x^n e^{-\lambda x}\,dx = \frac{1}{\lambda^{n+1}}\int_0^\infty t^n e^{-t}\,dt = \frac{\Gamma(n+1)}{\lambda^{n+1}}.$$

Therefore

$$\mathbb{E}[X] = \frac{\lambda^n}{\Gamma(n)} \cdot \frac{\Gamma(n+1)}{\lambda^{n+1}} = \frac{\Gamma(n+1)}{\Gamma(n)} \cdot \frac{1}{\lambda} = \frac{n}{\lambda}$$

Similarly, for the second moment:

$$\mathbb{E}[X^2] = \frac{\lambda^n}{\Gamma(n)} \int_0^\infty x^{n+1} e^{-\lambda x}\, dx = \frac{\lambda^n}{\Gamma(n)} \cdot \frac{\Gamma(n+2)}{\lambda^{n+2}} = \frac{\Gamma(n+2)}{\Gamma(n)} \cdot \frac{1}{\lambda^2} = \frac{n(n+1)}{\lambda^2}$$

Hence

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{n(n+1)}{\lambda^2} - \left(\frac{n}{\lambda}\right)^2 = \boxed{\frac{n}{\lambda^2}}$$

(c) The mean and variance derived above can be written as:

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[Y_i] = \frac{n}{\lambda}, \qquad \mathrm{Var}(X) = \sum_{i=1}^{n} \mathrm{Var}(Y_i) = \frac{n}{\lambda^2}.$$

which gives us

$$\mathbb{E}[Y_i] = \frac{1}{\lambda}, \qquad \mathrm{Var}[Y_i] = \frac{1}{\lambda^2}$$

A potential candidate for Y_i could be Exponential RV, i.e.

$$X = \sum_{i=1}^{n} Y_i, \quad \text{where } Y_i \sim \mathrm{Exp}(\lambda)\,.$$

Note that we cannot be sure of this just by comparing expectations. That is why it is a possible random variable. The density is that of a Gamma distribution with shape n and rate λ:

$$X \sim \mathrm{Gamma}(n, \lambda).$$

A well-known property is that the sum of n i.i.d. exponential random variables with common rate λ is $\mathrm{Gamma}(n, \lambda)$. We can prove this later by MGFs, covered later in the course.

Q2: Tail CDF Summation for a Discrete-Like Continuous Variable Let X be a continuous random variable with PDF $f_X(x) = e^{-x}$ for $x > 0$ (Standard Exponential).

Define a new discrete random variable $N = \lceil X \rceil$, which is the smallest integer greater than or equal to X.

(a) Find the probability mass function (PMF) of N, $P(N = n)$ for $n = 1, 2, 3, \ldots$.

(b) Calculate the expected value of N, $E[N]$, using the PMF.

(c) Use the "tail CDF summation" formula $E[N] = \sum_{n=0}^{\infty} P(N > n)$ to calculate $E[N]$.

A:

(a) $N = n$ means that $n - 1 < X \le n$.
Therefore:

$$\begin{aligned} P(N = n) = P(n - 1 < X \le n) &= \int_{n-1}^{n} f_X(x)\,dx && (1) \\ &= \int_{n-1}^{n} e^{-x}\,dx = [-e^{-x}]_{n-1}^{n} && (2) \\ &= -e^{-n} - (-e^{-(n-1)}) = e^{-(n-1)} - e^{-n} && (3) \\ &= e^{-n}(e - 1) && (4) \end{aligned}$$

Therefore: $\boxed{P(N = n) = (e - 1)e^{-n} \text{ for } n = 1, 2, 3, \ldots}$

(b)

$$\begin{aligned} E[N] &= \sum_{n=1}^{\infty} nP(N = n) = \sum_{n=1}^{\infty} n(e - 1)e^{-n} && (5) \\ &= (e - 1)\sum_{n=1}^{\infty} ne^{-n} = (e - 1)\sum_{n=1}^{\infty} n\left(\frac{1}{e}\right)^{n} && (6) \end{aligned}$$

Using the formula $\sum_{n=1}^{\infty} nr^n = \frac{r}{(1-r)^2}$ for $|r| < 1$:

$$\sum_{n=1}^{\infty} n\left(\frac{1}{e}\right)^{n} = \frac{1/e}{(1 - 1/e)^2} = \frac{1/e}{((e - 1)/e)^2} = \frac{e}{(e - 1)^2} \qquad (7)$$

Therefore:

$$\boxed{E[N] = (e - 1) \cdot \frac{e}{(e - 1)^2} = \frac{e}{e - 1}}$$

(c) First, we calculate $P(N > n)$:

$$\begin{aligned} P(N > n) &= P(\lceil X \rceil > n) = P(X > n) && (8) \\ &= \int_{n}^{\infty} e^{-x}\,dx = [-e^{-x}]_{n}^{\infty} = e^{-n} && (9) \end{aligned}$$

Now, using the tail summation formula:

$$\begin{aligned} E[N] &= \sum_{n=0}^{\infty} P(N > n) = \sum_{n=0}^{\infty} e^{-n} && (10) \\ &= \frac{1}{1 - e^{-1}} = \frac{1}{(e - 1)/e} = \frac{e}{e - 1} && (11) \end{aligned}$$

This confirms our previous result: $\boxed{E[N] = \frac{e}{e - 1}}$

Q3: Tail-Integral Formula

Let $X \ge 0$ be a continuous random variable with PDF $f_X(x)$ and let $p > 0$ with $\mathbb{E}[X^p] < \infty$. Derive the tail-integral identity:

$$\boxed{\mathbb{E}[X^p] \;=\; \int_0^{\infty} p\,s^{p-1}\,\mathbb{P}(X \ge s)\,ds}$$

A: Start with the definition of the p-th moment:

$$\mathbb{E}[X^p] = \int_0^\infty x^p \, f_X(x) \, dx.$$

For any $x \geq 0$ and $p > 0$, use the identity:

$$x^p = \int_0^x p \, s^{p-1} \, ds.$$

Substituting this into the expectation gives a double integral:

$$\mathbb{E}[X^p] = \int_0^\infty \left(\int_0^x p \, s^{p-1} \, ds \right) f_X(x) \, dx = \int_0^\infty \int_0^x p \, s^{p-1} f_X(x) \, ds \, dx.$$

Since the integrand is non-negative, Tonelli's theorem allows swapping the order of integration. When swapping, s goes from 0 to ∞, and for each fixed s, x runs from s to ∞:

$$\mathbb{E}[X^p] = \int_0^\infty p \, s^{p-1} \left(\int_s^\infty f_X(x) \, dx \right) ds.$$

Recognize the inner integral as the tail probability $\mathbb{P}(X \geq s)$:

$$\boxed{\mathbb{E}[X^p] = \int_0^\infty p \, s^{p-1} \, \mathbb{P}(X \geq s) \, ds}$$

The interchange is justified since the integrand is non-negative.

Q4: Gaussian Sensor Noise

A sensor records $X \sim N(\mu, \sigma^2)$. Measurements are accepted if they lie in $[\mu - \sigma, \mu + \sigma]$.

Find the probability that a measurement is accepted. You may use a standard normal table or calculator.

Solution:
Measurements are accepted if $X \in [\mu - \sigma, \mu + \sigma]$.

let $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$. Then

$$P(\mu - \sigma \leq X \leq \mu + \sigma) = P\left(\frac{\mu - \sigma - \mu}{\sigma} \leq Z \leq \frac{\mu + \sigma - \mu}{\sigma} \right).$$

That is,

$$P(-1 \leq Z \leq 1) = F_Z(1) - F_Z(-1),$$

where F_Z is the standard normal CDF.
By symmetry,

$$F_Z(1) - F_Z(-1) = 2F_Z(1) - 1.$$

Using the standard table value $F_Z(1) \approx 0.8413$,

$$2 \cdot 0.8413 - 1 = 0.6826.$$

So about 68.26% of sensor readings are accepted.

Q5: Let $X_1, X_2, \ldots, X_n$ be independent random variables, where X_j represents the time that a particular student takes to complete an exam. Assume that

$$X_j \sim \text{Exp}(\lambda), \quad j = 1, 2, \ldots, n.$$

(a) Find the distribution of

$$L = \min(X_1, X_2, \ldots, X_n),$$

the time until the *first* student completes the exam.

(b) Find the distribution of

$$T = \max(X_1, X_2, \ldots, X_n),$$

the time until *all* students have completed the exam.

(c) Specialize to the case $n = 3$. What is the expected value of T, i.e., the expected time at which all three students have completed the exam? Then extend the pattern to general n.

A:

(a) Since the X_j are independent,

$$\mathbb{P}(L > t) = \mathbb{P}(X_1 > t, \ldots, X_n > t) = \prod_{j=1}^{n} \mathbb{P}(X_j > t) = (e^{-\lambda t})^n = e^{-n\lambda t}.$$

Thus,

$$L \sim \text{Exp}(n\lambda).$$

(b) The maximum $T = \max(X_1, \ldots, X_n)$ has distribution function

$$\mathbb{P}(T \leq t) = \mathbb{P}(X_1 \leq t, \ldots, X_n \leq t) = \prod_{j=1}^{n} \mathbb{P}(X_j \leq t) = \left(1 - e^{-\lambda t}\right)^n, \quad t \geq 0.$$

Differentiating, the probability density function (PDF) is

$$p_T(t) = \frac{d}{dt}\left(1 - e^{-\lambda t}\right)^n = n\lambda e^{-\lambda t}\left(1 - e^{-\lambda t}\right)^{n-1}, \quad t \geq 0.$$

(c) For $n = 3$, let $T = \max(X_1, X_2, X_3)$. (The max does not have exponential distribution). We could integrate to get the expected value, but a better method is to use a decomposition based on order statistics and the memoryless property.
Write

$$T = T_1 + T_2 + T_3,$$

where

- $T_1 = \min(X_1, X_2, X_3) \sim \text{Exp}(3\lambda)$,
- T_2 is the additional time for the second student to finish. By memorylessness, after the first finish, the two remaining times are i.i.d. $\text{Exp}(\lambda)$, so $T_2 \sim \text{Exp}(2\lambda)$,

- T_3 is the additional time for the last student to finish, so $T_3 \sim \text{Exp}(\lambda)$.

Moreover, T_1, T_2, T_3 are independent by the memoryless property. Hence,

$$E[T] = E[T_1] + E[T_2] + E[T_3] = \frac{1}{3\lambda} + \frac{1}{2\lambda} + \frac{1}{\lambda} = \frac{11}{6\lambda}.$$

Extension to general n. The same argument yields the decomposition

$$T = \sum_{k=1}^{n} T_k, \quad \text{where } T_k \sim \text{Exp}((n-k+1)\lambda) \text{ independently,}$$

so

$$E[T] = \sum_{k=1}^{n} \frac{1}{(n-k+1)\lambda} = \frac{1}{\lambda} \sum_{m=1}^{n} \frac{1}{m} = \frac{H_n}{\lambda},$$

where H_n is the n-th harmonic number.

Q6: **Generalizing Transformations to Non-monotonic functions**

Let $Y = g(X)$, where g is a real-valued differentiable function and X is a continuous random variable with density $f_X(x)$. Denote the real roots of the equation

$$y = g(x)$$

by $x_1, x_2, \dots, x_n, \dots$. Show that the probability density function of Y is given by

$$f_Y(y) = \sum_i \frac{f_X(x_i)}{|g'(x_i)|},$$

where the sum is over all roots x_i for a given y, and $g'(x)$ denotes the derivative of $g(x)$.

A: We are given that $Y = g(X)$, where g is a real-valued differentiable function and X is a continuous random variable with density $f_X(x)$. The equation $y = g(x)$ has real roots $x_1, x_2, \dots, x_n, \dots$.

Step 1: Expression for $f_Y(y)$

We need to find $f_Y(y)$. By definition,

$$f_Y(y)\, dy = P(y < Y \le y + dy)$$

This probability is equal to the probability that X lies in the set of values such that $y < g(x) \le y + dy$.

From the figure and the fact that $g(x)$ is non-monotonic with multiple roots, this set consists of small intervals around each root x_i:

$$x_1 < x < x_1 + dx_1, \quad x_2 + dx_2 < x < x_2, \quad x_3 < x < x_3 + dx_3, \dots$$

where dx_i are small increments satisfying $dx_i = \frac{dy}{g'(x_i)}$, with appropriate sign consideration.

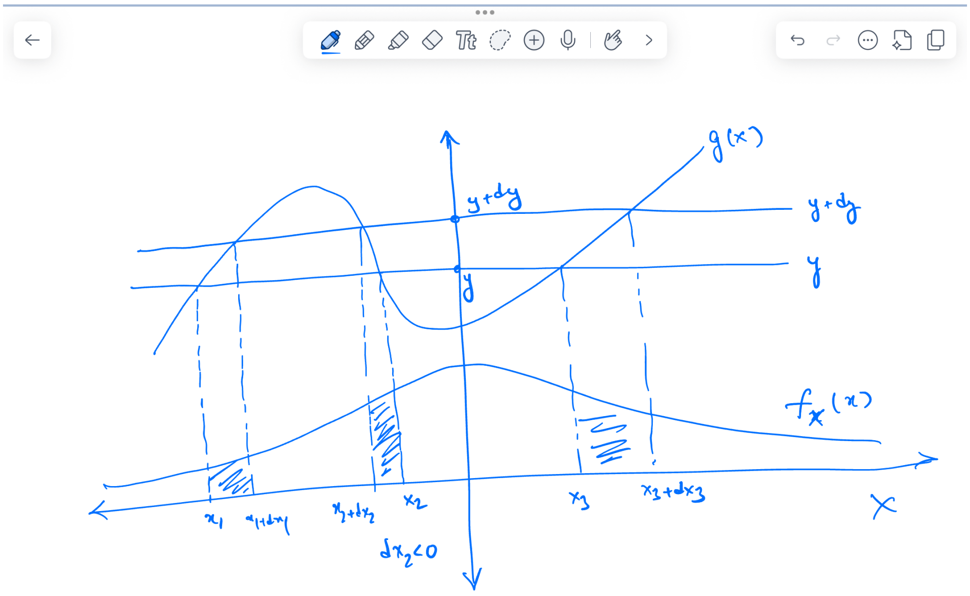

Figure 1: Illustration of the transformation $Y = g(X)$.

Thus,

$$\begin{aligned} P(y < Y \leq y + dy) &= P(x_1 < x < x_1 + dx_1) + P(x_2 + dx_2 < x < x_2) + P(x_3 < x < x_3 + dx \\ &= f_X(x_1)dx_1 + f_X(x_2)|dx_2| + f_X(x_3)dx_3 + \dots \end{aligned}$$

Using $dx_i = \frac{dy}{g'(x_i)}$, we get

$$f_Y(y)\, dy = \left(\frac{f_X(x_1)}{|g'(x_1)|} + \frac{f_X(x_2)}{|g'(x_2)|} + \frac{f_X(x_3)}{|g'(x_3)|} + \dots \right) dy$$

Step 2: Final Expression
Cancelling dy from both sides, we obtain

$$f_Y(y) = \sum_i \frac{f_X(x_i)}{|g'(x_i)|},$$

where the sum is over all roots x_i of $y = g(x)$.
Note: This assumes countable roots. For functions of bounded variation, uncountable roots can occur only if $g(x)$ is constant over some interval. If $g(x)$ is constant over some interval, say $g(x) = y_1$ for $x \in (x_a, x_b)$, then $F_Y(y)$ has a discontinuity at $y = y_1$. There would be a jump of $F_X(x_b) - F_X(x_a)$ and $f_Y(y)$ would be undefined for that region.

Q7: Mixed waiting time: taxi and bus [Shreyas]

A taxi stand and a bus stop are at the same location. When Shubham arrives there is a taxi already waiting with probability 1/2 (in which case he boards immediately). If there is no taxi waiting (probability 1/2), then the next taxi arrival time T (measured in minutes from now) is distributed uniformly on $[0, 8]$, while the next bus will arrive exactly in 5 minutes. Shubham takes whichever (taxi or bus) comes first. Let X denote Shubham's waiting time (in minutes).

(a) Find the CDF $F_X(x)$ for all $x \geq 0$, identify any point masses, and write down the PDF / mixed representation of X.
(b) Compute $\mathbb{E}[X]$.

A:
Since there is a taxi immediately with probability 1/2, we have a point mass at 0:

$$P(X = 0) = \tfrac{1}{2}.$$

Condition on the event that there is no taxi initially (call this event A^c), which has probability 1/2. In that case the waiting time equals $\min(T, 5)$ where $T \sim \text{Unif}(0, 8)$.
For $0 < x < 5$,

$$P(X \leq x) = P(X = 0) + P(A^c)\, P(T \leq x) = \tfrac{1}{2} + \tfrac{1}{2} \cdot \frac{x}{8} = \tfrac{1}{2} + \frac{x}{16}.$$

For $x \geq 5$,

$$P(X \leq x) = 1 - P(A^c)\, P(T > 5) = 1 - \tfrac{1}{2} \cdot \frac{3}{8} = 1 - \frac{3}{16} = \frac{13}{16},$$

but this equals $P(X \leq 5)$. For $x \geq 5$ the CDF reaches 1 after accounting the point mass at 5 (see remark below). Writing piecewise:

$$F_X(x) = P(X \leq x) = \begin{cases} 0, & x < 0, \\ \tfrac{1}{2} + \dfrac{x}{16}, & 0 \leq x < 5, \\ 1 - \dfrac{1}{2} P(T > 5) = 1 - \tfrac{1}{2} \cdot \dfrac{3}{8} = \dfrac{13}{16}, & x = 5, \\ 1, & x > 5. \end{cases}$$

Interpretation and point masses:

- $P(X = 0) = \tfrac{1}{2}$ (taxi already there).
- For $0 < x < 5$ there is a continuous density coming from the event "no initial taxi and taxi arrives before 5": differentiate F_X on $(0, 5)$ to get
$$f_X(x) = \frac{d}{dx} F_X(x) = \frac{1}{16}, \qquad 0 < x < 5.$$
- There is a point mass at $x = 5$ corresponding to the event "no initial taxi and taxi arrives after 5 (so bus at 5 is taken)": its probability is
$$P(X = 5) = P(A^c)\, P(T > 5) = \tfrac{1}{2} \cdot \frac{3}{8} = \frac{3}{16}.$$

Check total probability:

$$P(X=0)+\int_0^5 f_X(x)\,dx+P(X=5)=\tfrac{1}{2}+\frac{1}{16}\cdot 5+\tfrac{3}{16}=\tfrac{1}{2}+\tfrac{5}{16}+\tfrac{3}{16}=1.$$

(b) Expected waiting time.
Using the mixed representation,

$$\mathbb{E}[X]=0\cdot P(X=0)+\int_0^5 x\cdot\frac{1}{16}\,dx+5\cdot P(X=5).$$

Compute the integral:

$$\int_0^5 x\,\frac{1}{16}\,dx=\frac{1}{16}\cdot\frac{5^2}{2}=\frac{25}{32}.$$

And $5\cdot P(X=5)=5\cdot\frac{3}{16}=\frac{15}{16}=\frac{30}{32}$.
Hence

$$\mathbb{E}[X]=\frac{25}{32}+\frac{30}{32}=\frac{55}{32}\approx 1.71875$$

minutes.

Alternatively, condition on whether there is an initial taxi:

$$\mathbb{E}[X]=P(A)\cdot 0+P(A^c)\,\mathbb{E}[\min(T,5)]=\tfrac{1}{2}\,\mathbb{E}[\min(T,5)],$$

and for $T\sim\text{Unif}(0,8)$,

$$\mathbb{E}[\min(T,5)]=\int_0^5 t\cdot\frac{1}{8}\,dt+5\cdot P(T>5)=\frac{25}{16}+5\cdot\frac{3}{8}=\frac{55}{16},$$

so $\mathbb{E}[X]=\frac{1}{2}\cdot\frac{55}{16}=\frac{55}{32}$, agreeing with the previous calculation.

Q8: Start with an initial score of 0. In each turn, you generate a random number uniformly from (0,1) and add it to your score. What is the expected number of turns to reach a score > 2?

Hints:

(a) Try to model this in a recursive form, where you make a recursion based on what score you have, and how many turns do you need after seeing that score.

(b) Split this problem into 2 cases, when your sum is between $[0,1)$ and when it is between $[1,2)$.

(c) First solve the problem when you have sum ≥ 1, then use it to solve the case we are looking for.

A: Let $E(s)$ be the expected number of *additional* turns required to get a score > 2, given a current score of s. We want to find $E(0)$. The governing equation, found by conditioning on the first step, is:

$$E(s) = 1 + \int_0^1 E(s+x)\,dx$$

The boundary condition is $E(s) = 0$ for $s \geq 2$.

Case 1: $1 \leq s < 2$ We split the integral at the point where the new score $s + x$ crosses 2:

$$E(s) = 1 + \int_0^{2-s} E(s+x)\,dx + \int_{2-s}^1 \underbrace{E(s+x)}_{=0}\,dx = 1 + \int_s^2 E(u)\,du$$

Differentiating with respect to s gives $E'(s) = -E(s)$. The solution is $E(s) = Ce^{-s}$. Using the boundary condition $\lim_{s\to 2^-} E(s) = 1$, we find $Ce^{-2} = 1 \implies C = e^2$.

$$E(s) = e^{2-s} \quad \text{for } 1 \leq s < 2$$

Case 2: $0 \leq s < 1$ The full integral equation is $E(s) = 1 + \int_s^{s+1} E(u)\,du$. Differentiating gives:

$$E'(s) = E(s+1) - E(s)$$

For $s \in [0, 1)$, we know $s + 1 \in [1, 2)$, so we can use our previous result: $E(s + 1) = e^{2-(s+1)} = e^{1-s}$.

$$E'(s) + E(s) = e^{1-s}$$

This is a linear first-order DE. The integrating factor is e^s.

$$\frac{d}{ds}(e^s E(s)) = e^s(E'(s) + E(s)) = e^s e^{1-s} = e$$

Integrating gives $e^s E(s) = es + K$, so $E(s) = s + Ke^{-s}$. We use the boundary condition $E(1) = e^{2-1} = e$ to find K:

$$E(1) = 1 + Ke^{-1} = e \implies K = e(e-1) = e^2 - e$$

Thus, for $0 \leq s < 1$, the solution is $E(s) = s + (e^2 - e)e^{-s}$.
We need to find $E(0)$:

$$E(0) = 0 + (e^2 - e)e^{-0} = \mathbf{e^2 - e \approx 4.67078}$$