Homework 5

Asher Christian 006-150-286

2024-05-25

1 Problem 1

Suppose $X \sim f(x)$, and let $Y = aX + b$, where a and b are constants.

1. Prove $E(Y) = aE(X) + b$, and $Var(Y) = a^2 Var(X)$.

$$\begin{aligned}
E(Y) &= \int_{-\infty}^{\infty} Y(x) f(x) dx \\
&= \int_{-\infty}^{\infty} (aX + b) f(x) dx \\
&= a \int_{-\infty}^{\infty} x f(x) dx + b \int_{-\infty}^{\infty} f(x) dx \\
&= aE(x) + b
\end{aligned}$$

$$\begin{aligned}
Var(Y) &= \int_{-\infty}^{\infty} (Y(x) - E(Y))^2 f(x) dx \\
&= \int_{-\infty}^{\infty} (aX + b - aE(x) - b)^2 f(x) dx \\
&= a^2 \int_{-\infty}^{\infty} (X - E(X))^2 f(x) dx \\
&= a^2 Var(X)
\end{aligned}$$

2. Assuming $a > 0$, calculate the density of $Y, g(y)$ $X = \frac{Y-b}{a}$ and $\frac{dX}{dY} = \frac{1}{a}$

$$g(y) = f(x) * \frac{dx}{dy} = \frac{f(\frac{y-b}{a})}{a}.$$

2 Problem 2

Suppose $Z \sim N(0, 1)$, i.e.,

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

1. Calculate, $E(Z), Var(Z), E(|Z|)$

$$
\begin{aligned}
E(Z) &= \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
&= -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{u} du \\
&= -\frac{1}{\sqrt{2\pi}} \; e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} \\
&= -\frac{1}{\sqrt{2\pi}} + \frac{1}{\sqrt{2\pi}} \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
Var(Z) &= E(Z^2) - E(Z)^2 \\
&= E(Z^2) \\
E(Z^2) &= \int_{-\infty}^{\infty} \frac{z^2}{2\pi} e^{-\frac{z^2}{2}} dx \\
&= \frac{2}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2} dx \\
&= \frac{4}{\sqrt{\pi}} \int_{0}^{\infty} x^2 e^{-x^2} dx \\
&= \frac{2}{\sqrt{\pi}} \int_{0}^{\infty} z^{\frac{3}{2}-1} e^{-z} dz \\
&= \frac{2}{\sqrt{\pi}} \frac{\sqrt{\pi}}{2} \\
&= 1
\end{aligned}
$$

$$
\begin{aligned}
E(|Z|) &= \int_{-\infty}^{\infty} \frac{|x|}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
&= \frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} x e^{-\frac{x^2}{2}} dx \\
&= -\frac{\sqrt{2}}{\sqrt{\pi}} \; e^{-\frac{x^2}{2}} \Big|_{0}^{\infty} \\
&= 0 + \frac{\sqrt{2}}{\sqrt{\pi}} \\
&= \frac{\sqrt{2}}{\sqrt{\pi}}
\end{aligned}
$$

2. Let $X = \mu + \sigma Z$. Calculate the density of $X g(x), E(X), Var(X)$ based on

Problem 1.

$$
\begin{aligned}
g(x) &= \frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{(\frac{x-\mu}{\sigma})^2}{2}}}{\sigma} \\
&= \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}
\end{aligned}
$$

$$
\begin{aligned}
E(X) &= \sigma E(Z) + \mu \\
&= \mu
\end{aligned}
$$

$$
\begin{aligned}
Var(X) &= \sigma^2 Var(Z) \\
&= \sigma^2
\end{aligned}
$$

3. Suppose $P(Z \in [-2, 2]) = 95\%$ then what is $P(X \in [\mu - 2\sigma, \mu + 2\sigma])$?

$$
\begin{aligned}
95\% &= \int_{-2}^{2} \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dx \\
z &= \frac{x-\mu}{\sigma} \\
dz &= \frac{1}{\sigma} \\
x_1 &= \mu + 2\sigma \\
x_0 &= \mu - 2\sigma \\
95\% &= \int_{\mu-2\sigma}^{\mu+2\sigma} \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}dx
\end{aligned}
$$

3 Problem 3

Poisson process: Suppose we divide the time axis into small periods (0, Δt), (Δt, 2Δt), .., (t, t + Δt), ... Suppose within each interval, we flip a coin independently. Suppose the probability of getting a head is $\lambda\Delta$t. Let T be the time until the first head. Let X be the number of heads within [0, t]

1. Find the probability density function of T , and calculate $E(T)$ based on Geometric distribution.

$$
\begin{aligned}
P(T \in (t, t+\Delta t)) &= (1 - \lambda\Delta t)^{\frac{t}{\Delta t}}\lambda\Delta t \\
&= (e^{-\lambda\Delta t})^{\frac{t}{\Delta t}}\lambda\Delta t \\
&= e^{-\lambda t}\lambda\Delta t \\
&\sim e^{-\lambda t}\lambda dt
\end{aligned}
$$

$$\begin{aligned}
E(T) &= \int_0^\infty \lambda t e^{-\lambda t} dt \\
&= \frac{1}{\lambda} \int_0^\infty u e^{-u} du \\
w &= u \\
dw &= 1 \\
dv &= e^{-u} \\
v &= -e^{-u} \\
\int_0^\infty u e^{-u} du &= -u e^{-u} \Big|_0^\infty + \int_0^\infty e^{-u} du \\
&= 0 + -e^{-u} \Big|_0^\infty \\
&= 1 \\
E(T) &= 1 * \frac{1}{\lambda} \\
&= \frac{1}{\lambda}
\end{aligned}$$

2. Calculate $E(X)$ based on Binomial distribution. Find the probability mass function $P(X = k)$ as $\Delta t \to 0$.

$$t = n\Delta t.$$

$$n = \frac{t}{\Delta t}.$$

$$E(X) = np = \frac{t}{\Delta t}\lambda \Delta t = t\lambda.$$

$$\begin{aligned}
P(X \in (x, x + \Delta x)) &= \binom{n}{x} (\lambda \Delta t)^x (1 - \lambda \Delta t)^{n-x} \\
&= \frac{n!}{(n-x)!x!} (\lambda \Delta t)^x (1 - \lambda \Delta t)^{n-x} \\
&= \frac{\frac{t}{\Delta t}!}{(\frac{t}{\Delta t} - x)!x!} (\lambda \Delta t)^x (1 - \lambda \Delta t)^{\frac{t}{\Delta t} - x} \\
&= \frac{(\frac{t}{\Delta t})(\frac{t}{\Delta t} - 1)...(\frac{t}{\Delta t} - x + 1)}{x!} (\lambda \Delta t)^x (1 - \lambda \Delta t)^{\frac{t}{\Delta t}} (1 - \lambda \Delta t)^{-x} \\
&= \frac{(t)(t - \Delta t)(t - 2\Delta t)...(t - (x-1)\Delta t)}{x!} (\lambda)^x (1 - \lambda \Delta t)^{\frac{t}{\Delta t}} (1 - \lambda \Delta t)^{-x} \\
&\to \frac{(\lambda t)^x}{x!} e^{-\lambda t}
\end{aligned}$$

the approximation is as Δt approaches 0 making many terms vanish

4 Problem 4

Brownian motion or diffusion: Suppose a particle starts from 0, and within each period, it moves forward or backward by Δx, each with probability 1/2. Let X_t be the position at time t (assuming t is a multiple of Δt). Suppose there are n periods within $[0, t]$, i.e., $\Delta t = t/n$. Then we can write

$$X_t = \sum_{i=1}^{n} \epsilon_i \Delta x.$$

where $P(\epsilon_i = 1) = P(\epsilon_i = -1) = \frac{1}{2}$ and Z_i are independent

1. Calculate $E(X_t)$ and $Var(X_t)$.
   Let $Y \sim \text{Binomial}(n, \frac{1}{2})$

$$X_t = (Y - (n - Y))\Delta x = (2Y - n)\Delta x = (2Y - \frac{t}{\Delta t})\Delta x.$$

$$\begin{aligned} E(X_t) &= (2E(Y) - \frac{t}{\Delta t})\Delta x \\ &= (\frac{2t}{2\Delta t} - \frac{t}{\Delta t})\Delta x \\ &= 0 \\ Var(X_t) &= 4(\Delta X)^2 Var(Y) \\ &= 4(\Delta x)^2 n(\frac{1}{4}) \\ &= n(\Delta x)^2 \\ &= \frac{t(\Delta x)^2}{\Delta t} \end{aligned}$$

2. What is the relationship between Δx and Δt so that $Var(X_t)$ does not depend on discretization?

$$\frac{(\Delta x)^2}{\Delta t} = \sigma^2.$$

   σ^2 is a constant

$$\Delta x = \sigma\sqrt{\Delta t}.$$

3. According to the central limit theorem, what is the distribution of X_t?
   According to the central limit theorem, as t increases, the distribution approaches a normal distribution centered around 0 with Variance $\sigma^2 t$

5 Problem 5

1. Suppose we flip a fair coin 100 times independently. Let X be the number of heads. Based on normal approximation, find the 95% probability interval for X.

$$X \sim \text{Binomial}(n, \frac{1}{2}), n = 100.$$

$$\begin{aligned}
\mu &= \frac{n}{2} \\
\sigma^2 &= \frac{n}{4} \\
\sigma &= \frac{\sqrt{n}}{2}
\end{aligned}$$

Let $Z = \frac{X-\mu}{\sigma} = \frac{X-\frac{n}{2}}{\frac{\sqrt{n}}{2}}$
and $P(Z \in (a, b)) \to \int_a^b f(z)dz$ with $f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$

$$\begin{aligned}
95\% &= \int_{-a}^{a} f(z)dz \\
&= \left. \frac{\text{erf}(\frac{x}{\sqrt{2}})}{2} \right|_{-a}^{a} \\
1.9 &= \text{erf}(\frac{a}{\sqrt{2}}) - \text{erf}(-\frac{a}{\sqrt{2}}) \\
a &\approx 2 \\
\pm 2 &= \frac{x - \frac{n}{2}}{\frac{\sqrt{n}}{2}} \\
x &= \pm\sqrt{n} + \frac{n}{2} = 60, 40
\end{aligned}$$

To Verify

$$P(X \in (40, 60)) = \sum_{i=40}^{60} \frac{\binom{100}{i}}{2^{100}} \approx 0.964799799782.$$

2. Suppose 20% of the population support a candidate A. Suppose we randomly sample 100 people for the population (with replacement). Let $\hat{p} = X/100$ be the proportion of people in the sample who support candidate A. Based on normal approximation, find the 95% probability interval for $\hat{p}$.

$$X \sim \text{Bimomial}(100, \frac{1}{5}), \mu = 20, \sigma^2 = 16.$$

$$\hat{p} = \frac{X}{100}, \mu = 0.2, \sigma^2 = 0.0016, \sigma = 0.04.$$

$$Z = \frac{\hat{p} - 0.2}{0.04} \to P(Z \in (-2, 2)) \approx 95\%.$$
$$P(\hat{p} \in (0.12, 0.28)) = 95\%.$$

3. Suppose we randomly throw 10,000 points into the unit square $[0,1]^2$. Let A be the region $x^2 + y^2 \leq 1$. Let m be the number of points that fall into A. Let $\hat{\pi} = 4m/10000$ be our Monte Carlo estimate of π. What is the approximate normal distribution of $\hat{\pi}$? What is the 95% probability interval of $\hat{\pi}$?

$$m \sim \text{Bimomial}(10,000, \frac{\pi}{4}), \mu = \frac{10,000\pi}{4}, \sigma = 100\sqrt{\frac{\pi}{4}(1 - \frac{\pi}{4})}.$$
$$\hat{\pi} = \frac{4m}{10,000}, \mu = \pi, \sigma = \frac{\sqrt{\pi(1-\frac{\pi}{4})}}{50}.$$
$$\pm 2 = \frac{x - \pi}{\frac{\sqrt{\pi(1-\frac{\pi}{4})}}{50}}.$$
$$x = \pm\frac{1}{25}\sqrt{\pi(1 - \frac{\pi}{4})} + \pi.$$
$$P(\hat{\pi} \in (-\frac{1}{25}\sqrt{\pi(1 - \frac{\pi}{4})} + \pi, \frac{1}{25}\sqrt{\pi(1 - \frac{\pi}{4})} + \pi)) \approx 95\%.$$

6 Problem 6

1. Negative binomial distribution

$$P(X = k) = \binom{k + r - 1}{k}(1 - p)^k p^r.$$

2. Hyper-Geometric

$$P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}.$$

3. zipf

$$P(x) = \frac{x^{-(\rho+1)}}{\zeta(\rho + 1)}.$$

4. Chi-square

$$f(x;k) = \begin{cases} \frac{x^{\frac{k}{2}-1}e^{-\frac{x}{2}}}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})}, & x > 0; \\ 0, & \text{otherwise.} \end{cases}.$$

5. student t

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

6. Cauchy

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

7. Gamma

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}.$$

8. Beta

$$\frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathrm{B}(\alpha, \beta)}.$$

$$\mathrm{B}(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

9. Weibull

$$f(x; \lambda, k) = \begin{cases} \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}, & x \geq 0, \\ 0, & x < 0, \end{cases}.$$

10. Gumbel

$$\frac{1}{\beta} e^{-(z+e^{-z})}.$$

$$z = \frac{x - \mu}{\beta}.$$

11. Pareto

$$f_X(x) = \begin{cases} \frac{\alpha x_\mathrm{m}^\alpha}{x^{\alpha+1}} & x \geq x_\mathrm{m}, \\ 0 & x < x_\mathrm{m}. \end{cases}.$$