Chapter 4. Continuous Distributions

4.1 Probability density.

For all continuous variables, the probability mass function (pmf) is always equal to zero,

$$p(x) = 0 \quad \text{for all } x.$$

Cumulative distribution function (cdf):

$$F(x) = P\{X \leq x\} = P\{X < x\}.$$

Probability density function (pdf) $f(x)$:

$$f(x) = F'(x), \quad \text{is the derivative of the cdf.}$$

Then we have:

$$\int_a^b f(x)\,dx = F(b) - F(a) = P\{a < X < b\}$$

$b \leq X \leq b$

$a < X \leq b$

$a \leq X < b$

$$F(b) = P\{-\infty < X < b\} = \int_{-\infty}^{b} f(x)\,dx.$$

$$\int_{-\infty}^{+\infty} f(x)\,dx = P\{-\infty < X < +\infty\} = 1$$

Joint and marginal densities:

Joint cumulative distribution function:

$$F_{(X,Y)}(x,y) = P\{X \leq x \wedge Y \leq y\}$$

The joint density is the mixed derivative of the joint cdf.

$$f_{(X,Y)}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{(X,Y)}(x,y)$$

$$\begin{cases} f(x) = \int f(x,y)\,dy. \\ f(y) = \int f(x,y)\,dx. \end{cases}$$

if $f(x,y) = f(x)\cdot f(y) \rightarrow X, Y$ are independent.

Expectation and variance:

$$\boxed{u = E(X) = \int x\cdot f(x)\,dx}.$$

$$Var(X) = E(X-u)^2$$

$$= \int (x-u)^2 f(x)\,dx = \int (x^2 - 2ux + u^2) f(x)\,dx$$

$$= \int x^2 f(x)\,dx - 2u\underbrace{\int x f(x)\,dx}_{u} + \underbrace{u^2\int f(x)\,dx}_{1}$$

$$= \int x^2 f(x)\,dx - u^2$$

$$Cov(X,Y) = \iint (xy) f(x,y)\,dx\,dy - u_x u_y$$

4.2. Families of continuous distributions

1. Uniform distribution: the probability is only determined by the length of the interval, but not by its location. $f(x) = \frac{1}{b-a}$

For any $h>0$ and $t\in[a, b-h]$, the probability

$$P\{t<X<t+h\} = \int_t^{t+h} \frac{1}{b-a}\,dx = \frac{h}{b-a}.$$

is independent of t.

2. Standard Uniform distribution: the uniform distribution with $a=0$, $b=1$

Therefore, $f(x) = 1$ for $0<x<1$.

If X is a uniform (a,b) random variable, then

$$Y = \frac{X-a}{b-a} \Rightarrow X = a + (b-a)Y.$$

shift

scale.

Check that $X\in(a,b)$ if and only if $Y\in(0,1)$.

Expectation and variance:

$$E(Y) = \int y f(y)\,dy = \int_0^1 y\,dy = \frac{1}{2}.$$

$$Var(Y) = E(Y^2) - E^2(Y) = \int_0^1 y^2 dy - \left(\frac{1}{2}\right)^2 = \frac{1}{12}.$$

$$E(X) = E\{a + (b-a)Y\} = a + (b-a)E(Y) = \frac{a+b}{2}$$

$$Var(X) = Var\{a + (b-a)Y\} = (b-a)^2 Var(Y) = \frac{(b-a)^2}{12}$$

$$f(x) = \frac{1}{b-a}. \quad a < x < b.$$

Exponential distribution: (Model waiting time).

$$f(x) = \lambda \cdot e^{-\lambda x} \quad \text{for } x > 0.$$

$$F(x) = \int_0^x f(t)\,dt = \int_0^x \lambda \cdot e^{-\lambda t} dt = 1 - e^{-\lambda x} \quad (x > 0)$$

$$E(X) = \int t f(t)\,dt = \int_0^{\infty} t \cdot \lambda \cdot e^{-\lambda t} dt = \frac{1}{\lambda}.$$

$$Var(X) = \int t^2 f(t)\,dt - E^2(X)$$

$$= \int_0^{\infty} t^2 \cdot \lambda \cdot e^{-\lambda t} dt - \left(\frac{1}{\lambda}\right)^2$$

$$= \frac{1}{\lambda^2}$$

λ is frequency parameter, the number of events per time unit

Memoryless Property: (geometric distribution has this property).

$$P\{T > t + x \mid T > t\} = P\{T > x\}.$$

given. $P\{A \mid B\}$.

Gamma distribution. (Total time of a multistage scheme).

When a certain procedure consists of α independent steps, and each step takes Exponential(λ) amount of time, then the total time has Gamma distribution with parameters α and r.

$$f(x) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \cdot e^{-\lambda x}, \quad x > 0.$$

when $\alpha = 1$, the Gamma distribution becomes exponential.

Gamma cdf has the form:

$$F(t) = \int_0^t f(x)\,dx = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} \int_0^t x^{\alpha-1} \cdot e^{-\lambda x}\,dx$$

$$E(X) = \frac{\alpha}{\lambda}$$

$$E(X^2) = \frac{(\alpha+1)\alpha}{\lambda^2}$$

$$Var(X) = E(X^2) - E^2(X) = \frac{\alpha}{\lambda^2}$$

α, is shape parameter.

λ, frequency parameter.

Gamma-Poisson formula:

for a Gamma(α, λ) variable T, and a poisson(λt) variable X,

$P\{T > t\} = P\{X < \alpha\}$,

$P\{T \le t\} = P\{X \geq \alpha\}$.

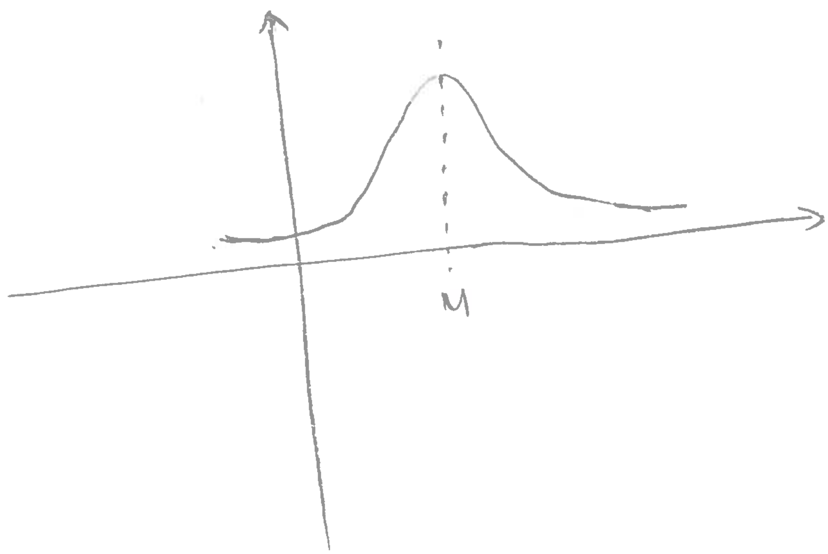
Normal distribution:

Normal distribution has a density:

$$f(x) = \frac{1}{\delta\sqrt{2\pi}} \exp\frac{-(x-u)^2}{2\delta^2} \quad , \quad -\infty < x < +\infty$$

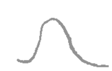
where $u = E(x)$, is the expectation. location parameter.

$\delta = Std(x) = \sqrt{Var(x)}$ → scale parameter.



bigger δ, flatter curves

smaller δ, narrower curves.



Standard Normal distribution: (P417 form)

when $u=0$, and $\delta=1$, the normal distribution is called standard Normal distribution.

We have:

z, standard normal random variable.

$\phi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-x^2}{2}}$, standard normal pdf.

$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} dx$, standard normal cdf.

Standardizing: for a non-standard normal random variable X with u, δ

$$Z = \frac{X-u}{\delta}$$

Unstandardizing: $X = u + \delta Z$

Symmetry :

$$P\{Z > x\} = P\{Z < -x\}.$$

$$\Phi(-z) = 1 - \Phi(z) \quad \text{for } -\infty < z < +\infty$$

4.3. Central Limit Theorem.

Let $X_1, X_2, \cdots$, be independent random variables with the same expectation $u = E(X_i)$ and the same standard deviation $\delta = \mathrm{Std}(X_i)$,

and Let $S_n = \sum_{i=1}^{n} X_i = X_1 + X_2 + \cdots + X_n$.

As $n \to \infty$, the Standardized sum

$$Z_n = \frac{S_n - E(S_n)}{\mathrm{Std}(S_n)}$$

Since $E(S_n) = n \cdot u$.

$$\mathrm{Std}(S_n) = \sqrt{\mathrm{Var}(S_n)} = \sqrt{n\delta^2} = \delta \cdot \sqrt{n}$$

Therefore, $Z_n = \dfrac{S_n - n \cdot u}{\delta \cdot \sqrt{n}}$

$$F_{Z_n}(z) = P\left\{\frac{S_n - n \cdot u}{\delta \cdot \sqrt{n}} \le z\right\} \to \Phi(z) \quad \text{for all } z.$$

For large n, $(n > 30)$. we can approx. sums using the standard normal distribution.

If $X_1, X_2, \cdots X_n$ are normally distributed, then S_n is normally distributed (works for small n)

For large n, we have: $P\{X = x\} = P\{x - 0.5 < X < x + 0.5\}$

continuity correction.

$$\text{Binomial}(n, p) \approx \text{Normal}(u = np, \delta = \sqrt{npq}).$$

Homework ch4 - Sec 3

Problem 1.

Among the watches manufactured by a particular company, 12% will be returned for warrenty repair. Use the Normal approximation to the Binomial distribution to find the probability that among 100 watches sold

a). 10 or fewer will be returned.

b). exactly 10 will be returned.

Solution: $p = 0.12$, $n = 100$, $q = 0.88$.

using normal approximation: $\mu = np = 12$, $\delta = \sqrt{npq} = \sqrt{100 \times 0.12 \times 0.88}$
$= 3.25$.

1). $P\{X \le 10\} = P\{-0.5 \le X \le 10.5\} = P\{\frac{-0.5-12}{3.25} < z < \frac{10.5-12}{3.25}\}$

$= P\{z < -0.46\} - P\{z < -3.85\}$

$= 0.3228 - 0.0001$

$= 0.3227$.

2). $P\{X = 10\} = P\{9.5 < X < 10.5\} = 0.1022$.

Problem 5.

Weight of elephants approximately follow an exponential distribut with the mean of 2.5 tons, one hundred elephants are being transported on a ship that has a cargo limit of 300 tons. what is the chance this ship will sink?

Solution: According to exponential distribution:

$E(X) = \mu = \frac{1}{\lambda} = 2.5$, $Var(X) = \frac{1}{\lambda^2}$, $\Rightarrow \delta = \frac{1}{\lambda} = \mu = 2.5$

$n = 100$. $S_n = 300$

$P\{X > 300\} = P\{z > \frac{300 - \mu n}{\sqrt{n} \cdot \delta}\} = P\{z > 2\} = 1 - P\{z < 2\}$
$= 0.0228$.