Review (3.4)

- Uniform distribution, its mean, variance, m.g.f.
- Exponential distribution and its mean, variance.
- Memoryless properties of exponential distribution.

Today's Lecture (3.7, 3.5)

- Comparison between discrete and continuous random variables
- Cauchy distribution, GOT A MOMENT?
- Gamma distribution and its connection to Poisson Process
- Mean, variance and m.g.f. of Gamma distribution.
- χ^2 distribution.

random variables

discrete

probability **mass** function

p.m.f.

$$p(x) = \mathrm{P}(\mathrm{X} = x)$$

$$\forall x \quad 0 \le p(x) \le 1$$

$$\sum_{\text{all } x} p(x) = 1$$

continuous

probability **density** function

p.d.f.

$$f(x)$$

$$\forall x \quad f(x) \ge 0$$

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

cumulative distribution function

c.d.f.

$$\mathrm{F}(x) = \mathrm{P}(\mathrm{X} \le x)$$

$$\mathrm{F}(x) = \sum_{y \le x} p(y)$$

$$\mathrm{F}(x) = \int_{-\infty}^{x} f(y)\,dy$$

expected value

$$\mathrm{E}(\mathrm{X}) = \mu_{\mathrm{X}}$$

discrete

If $\sum_{\text{all } x} |x| \cdot p(x) < \infty$,

$$\mathrm{E}(\mathrm{X}) = \sum_{\text{all } x} x \cdot p(x)$$

continuous

If $\int_{-\infty}^{\infty} |x| \cdot f(x)\,dx < \infty$,

$$\mathrm{E}(\mathrm{X}) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$$

discrete

If $\sum_{\text{all } x} |g(x)| \cdot p(x) < \infty$,

$$\mathrm{E}(g(\mathrm{X})) = \sum_{\text{all } x} g(x) \cdot p(x)$$

continuous

If $\int_{-\infty}^{\infty} |g(x)| \cdot f(x)\,dx < \infty$,

$$\mathrm{E}(g(\mathrm{X})) = \int_{-\infty}^{\infty} g(x) \cdot f(x)\,dx$$

variance

$$\mathrm{Var}(X) = \sigma_X^2 = E([X-\mu_X]^2) = E(X^2) - [E(X)]^2$$

discrete

$$\mathrm{Var}(X) = \sum_{\text{all } x} (x-\mu_X)^2 \cdot p(x)$$

$$= \sum_{\text{all } x} x^2 \cdot p(x) - [E(X)]^2$$

continuous

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x-\mu_X)^2 \cdot f(x)\,dx$$

$$= \left[\int_{-\infty}^{\infty} x^2 \cdot f(x)\,dx\right] - [E(X)]^2$$

moment-generating function

$$M_X(t) = E(e^{tX})$$

discrete

$$M_X(t) = \sum_{\text{all } x} e^{tx} \cdot p(x)$$

continuous

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} \cdot f(x)\,dx$$

<u>Example</u>

(Standard) Cauchy distribution: $f_X(x) = \dfrac{1}{\pi\left(1+x^2\right)}, \quad -\infty < x < \infty.$

Even though $f_X(x)$ is symmetric about zero, $\mathrm{E}(\mathrm{X})$ is undefined since

$$\int_{-\infty}^{\infty} |x| \cdot \frac{1}{\pi\left(1+x^2\right)}\, dx = \infty.$$

$$\mathrm{F}_X(x) = \int_{-\infty}^{x} \frac{1}{\pi\left(1+y^2\right)}\, dy = \frac{1}{\pi}\arctan(x) + \frac{1}{2}, \qquad -\infty < x < \infty.$$

$\mathrm{M}_X(0) = 1.$ $\qquad$ $\mathrm{M}_X(t)$ is undefined for all $t \neq 0$.

Volcano eruption (again)

Let N_t be the number of volcano eruptions to have occurred by time t, starting from now. Suppose that the volcano eruption forms a Poisson process with rate λ. Then $N_t \sim$ ______________.

Let X be the waiting time until the 4-th volcano eruption occurs and find the distribution of X.

$$P(X \le x) = P(N_x \geq 4)$$

$$F_X(x) = 1 - P(N_x \le 3)$$

$$= 1 - e^{-\lambda x}\left(\frac{(\lambda x)^0}{0!} + \frac{(\lambda x)^1}{1!} + \frac{(\lambda x)^2}{2!} + \frac{(\lambda x)^3}{3!}\right)$$

$$f_X(x) = F_X'(x) = -e^{-\lambda x}\cdot(-\lambda)\left[1 + \lambda x + \frac{(\lambda x)^2}{2} + \frac{(\lambda x)^3}{3!}\right]$$

$$- e^{-\lambda x}\cdot\left[\lambda + \lambda^2 x + \frac{\lambda^3 \cdot x^2}{2}\right]$$

$$= e^{-\lambda x}\cdot\lambda\,\frac{\lambda^3\cdot x^3}{3!} = e^{-\lambda x}\cdot\frac{\lambda^4\cdot x^3}{3!}$$

In general, if X is the waiting time until the α-th volcano eruption, then

$$F_X(x) = 1 - \sum_{k=0}^{\alpha-1} \frac{(\lambda x)^k e^{-\lambda x}}{k!}, \quad x > 0$$

and

$$f_X(x) = F_X'(x) = \frac{\lambda^\alpha x^{\alpha-1}}{(\alpha-1)!} e^{-\lambda x}, \quad x > 0$$

Gamma function

$$\Gamma(t) = \int_0^\infty y^{t-1} e^{-y} dy, \quad t > 0$$

Properties of $\Gamma(t)$:

- $\Gamma(t) = (t-1)\Gamma(t-1)$, $t > 1$.
- $\Gamma(1) = \int_0^\infty e^{-y} dy = 1$.
- When $t = n$, a positive integer,

$\Gamma(2) = 1! = 1$
$\Gamma(10) = 9!$

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

Gamma distribution

Definition: The random variable X has a gamma distribution if its p.d.f. is defined by

$$f(x) = \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} e^{-x/\theta}, \quad 0 \le x < \infty$$

Write Gamma(α, θ), where $\theta = 1/\lambda$, $\alpha > 0$ (not necessarily an integer). $\to \frac{1}{\lambda}$

See book P151 for plots

$$\begin{aligned}
M(t) &= \frac{1}{(1-\theta t)^\alpha}, \ t < 1/\theta \\
\mu &= M'(0) = \alpha\theta \\
\sigma^2 &= M''(0) - [M'(0)]^2 = \alpha(\alpha+1)\theta^2 - \alpha^2\theta^2 = \alpha\theta^2
\end{aligned}$$

Q: What is Exponential (θ)?

A: Gamma$(1, \theta)$. $(\alpha = 1)$

Example

Telephone calls enter a college switchboard at a mean rate of $1/2$ call per minute according to a Poisson process. Let X denote the waiting time until the second call arrives.

$\lambda = 1/2 \Rightarrow \theta = 2$

- What is the distribution of X?

Gamma $(2, 2)$

- What is the average waiting time?

$\mu = \alpha\theta = 4$

- What is the probability that the waiting time is longer than 3 minutes?

Method 1: $P(X>3) = 1 - F(3)$

$$= 1 - \left[1 - \sum_{k=0}^{1} \frac{(\frac{1}{2}x)^k e^{-\frac{1}{2}x}}{k!} \right] \Bigg|_{x=3}$$

$$= \sum_{k=0}^{1} \frac{(\frac{1}{2} \cdot 3)^k e^{-\frac{3}{2}}}{k!}$$

$$= e^{-\frac{3}{2}} + \frac{3}{2} \cdot e^{-\frac{3}{2}}$$

Method 2: N_3 : # of calls during time $[0, 3]$

then $N_3 \sim \text{Pois}(3\lambda) = \text{Pois}(\frac{3}{2})$

$$P(X>3) = P(N_3 < 2) = P(N_3 \le 1)$$

$$= e^{-\frac{3}{2}} \cdot \left(\frac{(\frac{3}{2})^0}{0!} + \frac{(\frac{3}{2})^1}{1!} \right)$$

$$= e^{-\frac{3}{2}} \left(1 + \frac{3}{2}\right)$$

Chi-square distribution

Gamma distribution:

$$f(x) = \frac{1}{\Gamma(\alpha)\theta^{\alpha}} x^{\alpha-1} e^{-x/\theta}, \quad 0 \leq x < \infty$$

Chi-square distribution: $\theta = 2$, $\alpha = r/2$, r is a positive integer.

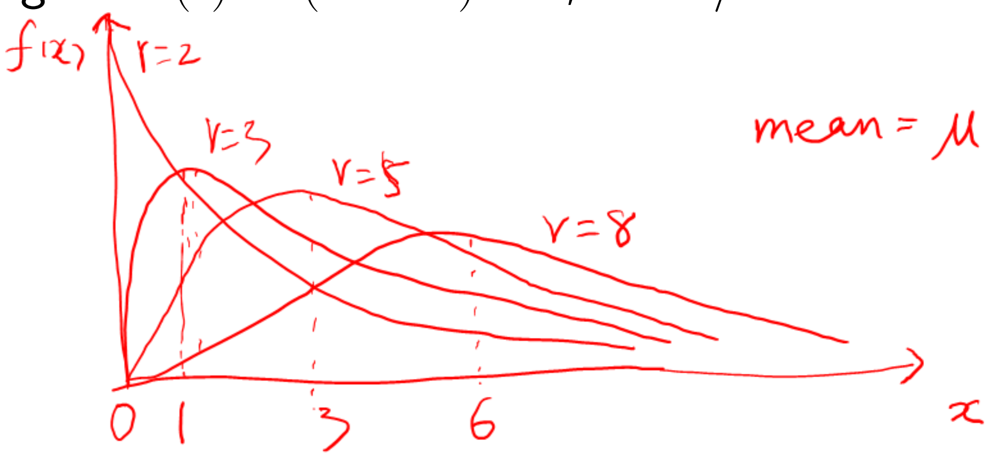
$$f(x) = \frac{1}{\Gamma(r/2)2^{r/2}} x^{r/2-1} e^{-x/2}, \quad 0 \leq x < \infty$$

r: degree of freedom. Write $X \sim \chi^2(r)$.

- Mean: $\mu = \alpha\theta = (r/2)2 = r$
- Variance: $\sigma^2 = \alpha\theta^2 = (r/2)2^2 = 2r$.
- m.g.f. $M(t) = (1 - 2t)^{-r/2}$, $t < 1/2$.

$$\chi^2(r) = \underbrace{N^2(0,1) + \cdots + N^2(0,1)}_{r \text{ of them}}$$

Plot:



mean $= \mu = \theta \cdot \alpha = r$