Chapter 3 Common Families of Distributions

3.1 Introduction

$N(\mu, \sigma^2)$

$N(0,1)$ $N(0, 0.1^2)$
$N(0, 0.5^2)$ ...
... $N(-1, 2^2)$

"family of distributions"

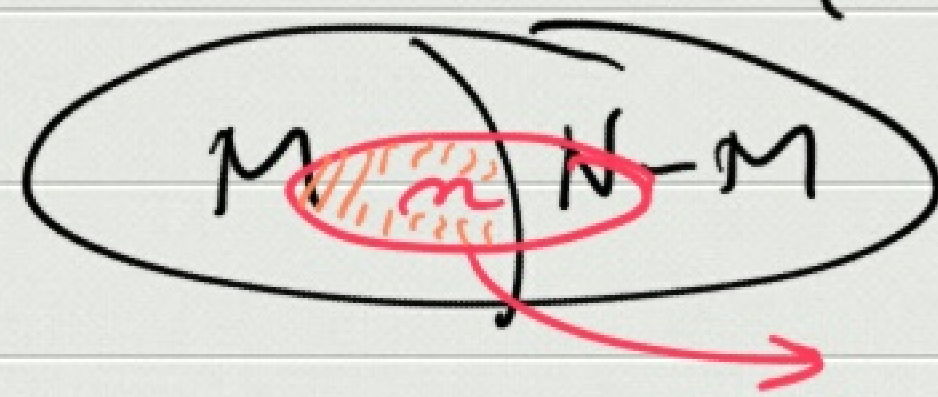
3.2 Discrete distributions

- X is said to have a discrete distⁿ if the range of X is countable.

Uniform: $P(X=x \mid N) = 1/N, \quad x = 1, 2, \cdots, N$

$E(X) = \frac{N+1}{2}, \quad \mathrm{Var}(X) = E(X^2) - \{E(X)\}^2 = \cdots.$

Hypergeometric:



X : # of 's

$$P(X=x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

Binomial:

$$X = \begin{cases} 1, & \text{w.p. } p \\ 0, & \text{w.p. } 1-p \end{cases} \qquad \text{Bernoulli}(p)$$

success

failure

$A_i = \{X=1 \text{ on the } i\text{th trial}\}$, $(i=1,2,\cdots,n)$, assumed indep.

Y = total # of successes in n trials :

$$P(Y=y \mid p) = \binom{n}{y} p^y (1-p)^{n-y}, \quad y=0,1,2,\cdots,n.$$

$$P[\underbrace{A_1 \cap A_2 \cap \cdots \cap A_y \cap A_{y+1}^c \cap A_{y+2}^c \cap \cdots \cap A_n^c}_{\text{shuffle} \cdots \Rightarrow \binom{n}{y}}] \overset{\text{independence}}{=} p^y (1-p)^{n-y}$$

$E(Y) = np$, $\quad Var(Y) = np(1-p)$

$M_Y(t) = [pe^t + (1-p)]^n$

Poisson

$$P(X=x \mid \lambda) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}, \quad x=0,1,\cdots$$

$$\left[e^{\lambda} = \frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots \quad (\text{Taylor series expansion}) \right]$$

$E(X) = \lambda$, $\quad Var(X) = \lambda$

$M_X(t) = e^{\lambda(e^t - 1)}$.

c.f. $B(n,p) \simeq \text{Poisson}(\lambda)$

when n is large and p is small in such a way $np \simeq \lambda$.

Negative binomial :

In a seq. of indep. Bernoulli(p) trials,

Let X denote the trial at which the r th success occurs. Then,

$$P(X=x \mid r,p) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad x=r, r+1, \cdots$$

- Or, Y = # of failures before the r-th success. Then $Y = X-r$.

$$P(Y=y) = \binom{r+y-1}{y} p^r (1-p)^y, \quad y=0,1,\cdots$$

$$\left(= (-1)^y \binom{-r}{y} p^r (1-p)^y \right)$$

: negative binomial.

$$E(Y) = r \cdot \frac{1-p}{p}, \quad Var(Y) = \frac{r(1-p)}{p^2}.$$

c.f. Poisson approx. for NB. (HW)

- See Example 3.2.6.

Geometric distⁿ: $P(X=x \mid p) = p(1-p)^{x-1}$, $x = 1, 2, \dots$ (NB with $r=1$)

$E(X) = 1/p$, $Var(X) = (1-p)/p^2$.

- Memoryless property: $\forall s>t$, $P(X>s \mid X>t) = P(X>s-t)$.

3.3 Continuous distⁿ

Uniform: $f(x \mid a,b) = \dfrac{1}{b-a} \cdot I(x \in [a,b])$

$E(X) = \dfrac{a+b}{2}$, $Var(X) = \dfrac{(b-a)^2}{12}$.

Gamma: $f(x \mid \alpha, \beta) = \dfrac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta} \cdot I(x>0)$

$E(X) = \alpha\beta$, $Var(X) = \alpha\beta^2$, $M_X(t) = (1-\beta t)^{-\alpha}$, $t < 1/\beta$.

• Gamma – Poisson relationship

$X \sim \mathrm{Gamma}(\alpha, \beta)$, α : an integer.

Then, $P(X \le x) \overset{\forall x}{=} P(Y \ge \alpha)$ where $Y \sim \mathrm{Poisson}(x/\beta)$.

(pf.) $$P(X \le x) = \int_0^x \frac{1}{\Gamma(\alpha)\beta^\alpha} t^{\alpha-1} e^{-t/\beta} dt \qquad \longrightarrow (-\beta e^{-t/\beta})'$$

$$= \frac{1}{(\alpha-1)!\,\beta^\alpha}\left[-t^{\alpha-1}\beta e^{-t/\beta}\Big|_0^x + \int_0^x (\alpha-1) t^{\alpha-2} \beta e^{-t/\beta}\, dt \right]$$

$$= \underbrace{- \frac{(t/\beta)^{\alpha-1} e^{-t/\beta}}{(\alpha-1)!}}_{P(Y=\alpha-1)} + \underbrace{\frac{1}{(\alpha-2)!\,\beta^{\alpha-1}} \int_0^x t^{\alpha-2} e^{-t/\beta} dt}_{P(\mathrm{Gamma}(\alpha-1, \beta) \le x)}$$

$$= \cdots$$

$$= -P(Y=\alpha-1) - P(Y=\alpha-2) - \cdots - P(Y=0) + 1 = P(Y \ge \alpha) \quad //$$

• Special cases of the gamma family

① $\alpha = p/2$ (p: an integer), $\beta = 2$: χ^2_p.

② $\alpha = 1$: exponential with scale parameter β.

"memoryless property": If $X \sim$ exponential(β), then $\forall\, s > t \geq 0$,

$$P(X > s \mid X > t) = P(X > s - t).$$

pf

$$P(X>s \mid X>t) = P(X>s,\, X>t) / P(X>t)$$
$$= P(X>s) / P(X>t)$$
$$= e^{-(s-t)/\beta}. \quad //$$

③ $X \sim$ exponential(β), then $Y = X^{1/\gamma} \sim$ Weibull(γ, β)

$$f_Y(y \mid \gamma, \beta) = \frac{\gamma}{\beta} y^{\gamma-1} e^{-y^{\gamma}/\beta}, \quad y > 0$$

Normal :

$X \sim N(\mu, \sigma^2)$, $f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $x \in \mathbb{R}$.

$\left| \begin{matrix} \frac{\partial t}{\partial r} & \frac{\partial t}{\partial \theta} \\ \frac{\partial s}{\partial r} & \frac{\partial s}{\partial \theta} \end{matrix} \right| = r$

· $Z = (X-\mu)/\sigma \sim N(0,1)$

· $\int_0^\infty e^{-z^2/2}\, dz \underset{(t=z^2)}{=} \int_0^\infty e^{-t/2} \cdot \frac{1}{2} t^{-1/2}\, dt = \frac{\Gamma(\frac{1}{2})}{\sqrt{2}} \int_0^\infty \text{Gamma}(\frac{1}{2}, 2)\, dt = \frac{1}{\sqrt{2}} \Gamma(\frac{1}{2})$

$$\left(\int_0^\infty e^{-z^2/2}\, dz\right)^2 = \int_0^\infty e^{-t^2/2}\, dt \cdot \int_0^\infty e^{-s^2/2}\, ds = \int_0^\infty\int_0^\infty e^{-(t^2+s^2)/2}\, dt\, ds$$

$\begin{cases} t = r\cos\theta \\ s = r\sin\theta \end{cases}$: $t^2+s^2 = r^2$, $dt\,ds = r\,d\theta\,dr$ with $0<\theta<\frac{\pi}{2}$, $r>0$.

$$= \int_0^\infty \int_0^{\pi/2} r e^{-r^2/2}\, d\theta\, dr = \frac{\pi}{2} \int_0^\infty r e^{-r^2/2}\, dr = \frac{\pi}{2}\left(-e^{-r^2/2}\right)\Big|_0^\infty = \frac{\pi}{2}$$

$\therefore \Gamma(\frac{1}{2}) = \sqrt{\pi}$.

- $B(n,p) \simeq N(np, np(1-p))$ when n is large and p is not extreme.

(e.g.) $X \sim B(25, .6)$

$$P(X \le 13) = \sum_{x=0}^{13} \binom{25}{x} (.6)^x (.4)^{25-x} = .267$$

$Y \sim N(15, 2.45^2)$, $25 \times 0.6 = 15$, $25 \times 0.6 \times 0.4 = 2.45^2$

$$P(X \le 13) = P(X \le 13.5) \simeq P(Y \le 13.5) = P(Z \le -0.61) = 0.271$$

Beta distribution: $f(x|\alpha,\beta) = \dfrac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}\,dx, \quad 0 \le x \le 1$

where $B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx$ is the beta func.

c.f. $B(\alpha,\beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$.

$E(X) = \dfrac{\alpha}{\alpha+\beta}$, $\quad Var(X) = \dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

Cauchy dist^n: $f(x|\theta) = \dfrac{1}{\pi} \cdot \dfrac{1}{1+(x-\theta)^2}, \quad x \in \mathbb{R}$

$E|X| = \infty$: no moments exist.

- What is it good for? : $N(0,1)/N(0,1) \sim$ Cauchy.

Lognormal dist^{n}: $$f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \frac{1}{x} \cdot e^{-(\log x - \mu)^2/(2\sigma^2)}, \quad x > 0$$

$Y = \log X \sim N(\mu, \sigma^2)$

$$E(X) = E\left(e^{\log X}\right) = E(Y) = e^{\mu + \sigma^2/2}, \quad Var(X) = e^{2(\mu+\sigma^2)} - e^{2\mu + \sigma^2}$$

Double exponential $$f(x|\mu,\sigma) = \frac{1}{2\sigma} e^{-|x-\mu|/\sigma}, \quad x \in \mathbb{R}$$

$E(X) = \mu, \quad Var(X) = 2\sigma^2.$

3.4 Exponential family

- A family of pdfs (or pmfs) is called an exponential family if it can be expressed as

$$f(x|\underset{\sim}{\theta}) = h(x)\, c(\underset{\sim}{\theta}) \exp\left(\sum_{i=1}^{k} w_i(\underset{\sim}{\theta})\, t_i(x)\right). \qquad (*)$$

(e.g.) $B(n,p)$: Fix n.

$$f(x|p) = \binom{n}{x} p^x (1-p)^{n-x} = \binom{n}{x}(1-p)^n \left(\frac{p}{1-p}\right)^x = \underbrace{\binom{n}{x}}_{h(x)} \underbrace{(1-p)^n}_{c(p)} \exp\left(\underbrace{\log\frac{p}{1-p}}_{w_1(\theta)} \cdot \underbrace{x}_{t_1(x)}\right)$$

Thm

$$E\left(\sum_{i=1}^{k} \frac{\partial w_i(\underset{\sim}{\theta})}{\partial \theta_j} t_i(X)\right) = -\frac{\partial}{\partial \theta_j} \log c(\underset{\sim}{\theta})$$

$$Var\left(\sum_{i=1}^{k} \frac{\partial w_i(\underset{\sim}{\theta})}{\partial \theta_j} t_i(X)\right) = -\frac{\partial^2}{\partial \theta_j^2} \log c(\underset{\sim}{\theta}) - E\left(\sum_{i=1}^{k} \frac{\partial^2 w_i(\underset{\sim}{\theta})}{\partial \theta_j^2} t_i(X)\right).$$

(e.g.) $B(n,p)$ (cont'd)

$$\left.\begin{aligned} \frac{d}{dp} w_1(p) &= \frac{1}{p(1-p)} \\ \frac{d}{dp} \log c(p) &= -\frac{n}{1-p} \end{aligned}\right\} \Rightarrow E\left(\frac{1}{p(1-p)} X\right) = \frac{n}{1-p} \qquad \therefore E(X) = np.$$

(e.g.) $N(\mu, \sigma^2)$, $\underset{\sim}{\theta} = (\mu, \sigma) \in (-\infty, \infty) \times (0, \infty)$

$$f(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) = \underbrace{1}_{h(x)} \cdot \underbrace{\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\mu^2}{2\sigma^2}\right)}_{c(\underset{\sim}{\theta}) = c(\mu,\sigma)} \exp\left(\underbrace{\frac{1}{\sigma^2}}_{w_1(\underset{\sim}{\theta})} \cdot \underbrace{\left(-\frac{x^2}{2}\right)}_{t_1(x)} + \underbrace{\frac{\mu}{\sigma^2}}_{w_2(\underset{\sim}{\theta})} \cdot \underbrace{x}_{t_2(x)}\right)$$

(e.g.) $f(x|\theta) = \frac{1}{\theta} \exp\left(1 - \frac{x}{\theta}\right) \underbrace{I(0 < \theta < x)}$

~~exp'l family~~

- An exp'l family is reparametrized as

$$f(x|\eta) = h(x)\, c^*(\underset{\sim}{\eta})\, \exp\left(\sum_{i=1}^{k} \eta_i t_i(x)\right) \qquad \text{(canonical form)}$$

$$\mathcal{H} = \left\{ \underset{\sim}{\eta} = (\eta_1, \cdots, \eta_k) \;\middle|\; \int_{-\infty}^{\infty} h(x) \exp\left(\sum_{i=1}^{k} \eta_i t_i(x)\right) dx < \infty \right\} \quad \text{(the natural parameter space)}$$

- $\{ (w_1(\underset{\sim}{\theta}), w_2(\underset{\sim}{\theta}), \cdots, w_k(\underset{\sim}{\theta})) \mid \underset{\sim}{\theta} \in \Theta \} \subset \mathcal{H}$.
- $\mathcal{H}$ is convex.

(e.g.) $N(\mu, \sigma^2)$ (cont'd)

$$f(x|\eta_1, \eta_2) = \underbrace{1}_{h(x)} \cdot \underbrace{\frac{\sqrt{\eta_1}}{\sqrt{2\pi}} \exp\left(-\frac{\eta_2^2}{2\eta_1}\right)}_{c^*(\eta_1, \eta_2)} \exp\left(\eta_1 \cdot \underbrace{\left(-\frac{x^2}{2}\right)}_{t_1(x)} + \eta_2 \cdot \underbrace{x}_{t_2(x)}\right)$$

$\eta_1 = 1/\sigma^2, \quad \eta_2 = \mu/\sigma^2$

$\mathcal{H} = (0, \infty) \times (-\infty, \infty)$.

Defn A curved exp'l family is a family of pdfs of the form (*) for which the dimension of $\underset{\sim}{\theta}$ is equal to $d<k$.

(e.g.) $N(\mu, \mu^2)$, $f(x|\mu) = \frac{1}{\sqrt{2\pi}\mu} \exp\left(-\frac{(x-\mu)^2}{2\mu^2}\right)$

$\underset{\sim}{\theta} = (\mu, \mu^2)$ //

(e.g.) $X_1, \cdots, X_n \overset{iid}{\sim} \text{Poisson}(\lambda)$

$\bar{X} \mathrel{\dot\sim} N(\lambda, \lambda/n)$, $\underset{\sim}{\theta} = (\lambda, \lambda/n)$ //

(e.g.) $X_1, \cdots, X_n \overset{iid}{\sim} B(n, p)$

$\bar{X} \mathrel{\dot\sim} N(p, p(1-p)/n)$, $\underset{\sim}{\theta} = (p, \frac{p(1-p)}{n})$ //

3.5 Location and scale families

<u>Thm</u>: $f(x)$, a pdf.

$\mu \in \mathbb{R}$, $\sigma > 0$: given constants

then $g(x|\mu,\sigma) = \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right)$ is a pdf.

<u>Defn</u>: $f(x)$, a pdf.

$\{ f(x-\mu) \mid \mu \in \mathbb{R} \}$ is called the location family with standard pdf $f(x)$ and μ is the location parameter for the family.

(e.g.) $\{ N(\mu, 1) \mid \mu \in \mathbb{R} \}$

$\{ \mathrm{Exp}(\mu) \mid \mu \in \mathbb{R} \}$ "$f(x|\mu) = e^{-(x-\mu)} 1(x-\mu > 0)$"

- If $X \sim f(x-\mu)$, then X may be represented as $X = Z + \mu$, where $Z \sim f(x)$. (will be proved later)

<u>Defn</u> : $f(x)$, a pdf.

$\{ \frac{1}{\sigma} f(\frac{x}{\sigma}) \mid \sigma > 0 \}$ is called the scale family with standard pdf $f(x)$ and σ is called the scale parameter of the family.

- If $X \sim \frac{1}{\sigma} f(\frac{x}{\sigma})$, then X may be represented as $X = \sigma Z$, $Z \sim f(x)$.

Defn: $f(x)$, a pdf.

$\left\{ \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right) \,\middle|\, \mu \in \mathbb{R}, \sigma > 0 \right\}$: the location-scale family with standard pdf $f(x)$

: μ is the location parameter and σ is the scale parameter.

Thm: $X \sim \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right)$. Then $X = \mu + \sigma Z$, $Z \sim f(x)$.

pf. (변수변환법)

HW! "중간고사 출제"

Thm: $Z \sim f(x)$, $E(Z)$ and $Var(Z)$ exist.

If $X \sim \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right)$, then $E(X) = \mu + \sigma E(Z)$ and $Var(X) = \sigma^2 Var(Z)$.

- For any location-scale family with a finite mean and variance, the standard pdf $f(x)$ can be chosen in such a way that

  $E(Z) = 0$ and $Var(Z) = 1$.

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

3.6 Inequalities and identities.

- Chebyshev's ineq.: for any $r > 0$,

$$P(g(X) \geq r) \leq \frac{E g(X)}{r}$$

pf:

$$P(g(X) \geq r) = \int_{\{x \mid g(x) \geq r\}} f_X(x)\, dx$$

$$\leq \int_{\{x \mid g(x) \geq r\}} \frac{1}{r} g(x) f_X(x)\, dx$$

$$\leq \frac{1}{r} \int_{-\infty}^{\infty} g(x) f_X(x)\, dx = \frac{1}{r} E g(X). \quad //$$