Review

Hyonho Chun

Department of Mathematical Sciences
KAIST

March 18, 2020

Common Families of Distributions

Please review the following pdfs.

- Discrete case:
  Discrete uniform, Hypergeometric, Binomial, Negative Binomial, Poisson, Geometric distributions.
- Continuous case:
  Uniform, Gamma, Normal, Beta, Cauchy, Lognormal, Double exponential distributions.

Exponential families

- A family of pdfs or pmfs is called an exponential family if it can be written as
  $f(x|\boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta})\exp\left(\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(x)\right)$
- Theorem 3.4.2
  If X is a random variable with pdf or pmf takes the form of the exponential family distribution then
  $E\left(\sum_{i=1}^{k} \frac{\partial w_i(\boldsymbol{\theta})}{\partial \theta_j} t(X)\right) = -\frac{\partial}{\partial \theta_j} \log c(\boldsymbol{\theta});$
  $Var\left(\sum_{i=1}^{k} \frac{\partial w_i(\boldsymbol{\theta})}{\partial \theta_j} t_i(X)\right) = -\frac{\partial^2}{\partial \theta_j^2} \log c(\boldsymbol{\theta}) - E\left(\sum_{i=1}^{k} \frac{\partial^2 w_i(\boldsymbol{\theta})}{\partial \theta_j^2} t_i(X)\right)$
- Reparametrized form
  $f(x|\eta) = h(x)c^{\star}(\eta)\exp\left(\sum_{i=1}^{k} \eta_i t_i(x)\right)$
  $\mathscr{H} = \left\{\eta = (\eta_1, \ldots, \eta_k) : \int_{-\infty}^{\infty} h(x) exp(\sum_{i=1}^{k} \eta_i t_i(x)) dx < \infty\right\}$: Natural parameter space.

Exponential families

- $1 = \int h(x)c(\boldsymbol{\theta}) \exp(\sum_{i=1}^k w_i(\boldsymbol{\theta}) t_i(x)) dx$
- $0 = \frac{d}{d\theta_j} \int h(x)c(\boldsymbol{\theta}) \exp(\sum_{i=1}^k w_i(\boldsymbol{\theta}) t_i(x)) dx$
- $0 = \int \frac{\partial}{\partial\theta_j} (h(x)c(\boldsymbol{\theta}) \exp(\sum_{i=1}^k w_i(\boldsymbol{\theta}) t_i(x))) dx$
- $0 = \int h(x)(\frac{\partial}{\partial\theta_j} c(\boldsymbol{\theta})) \exp(\sum_{i=1}^k w_i(\boldsymbol{\theta}) t_i(x)) dx +$
  $h(x)c(\boldsymbol{\theta})(\sum_{i=1}^k \frac{\partial w_i(\boldsymbol{\theta})}{\partial\theta_j} t_i(x)) \exp(\sum_{i=1}^k w_i(\boldsymbol{\theta}) t_i(x)) dx$
- Applying one more derivative will yield the second equality.

Bivariate transformation

- For one-to-one transformations, $(u,v) = (g_1(x,y), g_2(x,y))$
  Inverse mapping: $(x,y) = (h_1(u,v), h_2(u,v))$
  $J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$
  $f_{U,V}(u,v) = f_{X,Y}(h_1(u,v), h_2(u,v))|J|$
- Example
  $X \sim \text{beta}(\alpha, \beta)$ and $Y \sim \text{beta}(\alpha+\beta, \gamma)$: independent random variables.
  Let $U = XY$ and $V = X$. Find the joint pdf of (U,V) and the marginal pdf of U.

Bivariate transformation

- $0 \le u \le v \le 1$
- $x = v$ and $y = u/v$,
- $|J| = \begin{vmatrix} 0 & 1 \\ 1/v & -u/v^2 \end{vmatrix} = 1/v$
- $f_{U,V}(u,v) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} v^{\alpha-1}(1-v)^{\beta-1} \frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha+\beta)\Gamma(\gamma)} (u/v)^{\alpha+\beta-1}(1-u/v)^{\gamma-1}\frac{1}{v}$ when $0 \le u \le v \le 1$ and 0 otherwise.
- $f_U(u) = \int_u^1 f_{U,V}(u,v)dv$
- Using a special transformation of $t = \frac{\frac{u}{v}-u}{1-u}$, $f_U(u) = \frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta+\gamma)} u^{\alpha-1}(1-u)^{\beta+\gamma-1}$ when $0 < u < 1$.

Bivariate transformation

- $\mathscr{A} = \{(x,y) : f_{X,Y}(x,y) > 0\}$
- $U = g_1(X,Y)$, $V = g_2(X,Y)$ is one-to-one transformation from A_i onto $\mathscr{B}$ for each $i = 1, 2, \ldots, k$.
- Denote the ith inverse by $x = h_{1i}(u,v)$ and $y = h_{2i}(u,v)$.
  $f_{U,V}(u,v) = \sum_{i=1}^{k} f_{X,Y}(h_{1i}(u,v), h_{2i}(u,v))|J_i|$.
- Example: X, Y independent $N(0,1)$ random variables.
  $U = X/Y$ and $V = |Y|$.
  Find the pdf of (U,V) and U.

Bivariate transformation

- $f_{X,Y}(x,y) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2) \frac{1}{\sqrt{2\pi}} \exp(-y^2/2)$
- $A_1 = \{(x,y) : y > 0\}$, $A_2 = \{(x,y) : y < 0\}$
- On A_1, $x = uv$ and $y = v$
  $|J_1| = \left| \begin{pmatrix} v & u \\ 0 & 1 \end{pmatrix} \right| = v$
- Similarly, on A_2, $x = uv$ and $y = -v$
  $|J_2| = \left| \begin{pmatrix} v & u \\ 0 & -1 \end{pmatrix} \right| = v$
- $$\begin{aligned} f_{U,V}(u,v) &= \sum_{i=1}^{2} f_{X,Y}(h_{1i}(u,v), h_{2i}(u,v))|J_i| \\ &= \frac{1}{\pi} \exp(-(u^2v^2 + v^2)/2)v \end{aligned}$$
  when $v > 0$ and 0 otherwise.
- $f_U(u) = \int_0^\infty \frac{1}{\pi} exp(-(u^2+1)v^2/2)vdv = \frac{1}{\pi(1+u^2)}$, $-\infty < u < \infty$.

Multivariate transformation

- $f_X(x_1, x_2, x_3, x_4) = 24e^{-x_1-x_2-x_3-x_4}$, $0 < x_1 < x_2 < x_3 < x_4 < \infty$.
- Consider the transformation
  $U_1 = X_1,\ U_2 = X_2 - X_1,\ U_3 = X_3 - X_2,\ U_4 = X_4 - X_3$.
  Find the joint pdf of U.

Multivariate transformation

- $x_1 = u_1,\ x_2 = u_2 + u_1,\ x_3 = u_3 + u_2 + u_1,\ x_4 = u_4 + u_3 + u_2 + u_1$
- $J = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$
- $\mathscr{B} = \{(u_1, u_2, u_3, u_4) : u_1 > 0, u_2 > 0, u_3 > 0, u_4 > 0\}$
- $f_{\mathbf{U}}(u_1, u_2, u_3, u_4) = 24\exp(-4u_1 - 3u_2 - 2u_3 - u_4)$ on $\mathscr{B}$
- $f_{U_1}(u_1) = \int_0^\infty \int_0^\infty \int_0^\infty 24\exp(-4u_1 - 3u_2 - 2u_3 - u_4) du_2 du_3 du_4 = 4exp(-4u_1)$
- $f_{U_4}(u_4) =?$

Inequalities

- Chebychev's inequality
  X: a random variable, $g(x)$ is nonnegative function.
  $P(g(X) \geq r) \leq \frac{Eg(X)}{r}$
  $\langle$ Proof $\rangle$

$$
\begin{aligned}
Eg(X) &\geq \int_{\{g(x) \geq r\}} g(x) f_X(x) dx \\
&\geq r \int_{\{g(x) \geq r\}} f_X(x) dx \\
&= rP(g(X) \geq r)
\end{aligned}
$$

- (Example)
  $P(X > a) = P(e^{tX} > e^{ta}) \leq e^{-ta} M_X(t)$.

Inequalities

- (Lemma) Let a and b be any positive numbers and let p and q be any positive numbers satisfying
  $\frac{1}{p} + \frac{1}{q} = 1.$
  Then $\frac{1}{p}a^p + \frac{1}{q}b^q \geq ab$ with equality if and only if $a^p = b^q$.
- Theorem 4.7.2 (Holder's inequality)
  Let X and Y be any two random variables, and let p and q satisfy $\frac{1}{p} + \frac{1}{q} = 1.$ Then
  $|EXY| \leq E|XY| \leq (E|X|^p)^{1/p}(E|Y|^q)^{1/q}.$
  $\langle Proof \rangle$
  Setting $a = \frac{|X|}{(E|X|^p)^{1/p}}$ and $b = \frac{|Y|}{(E|Y|^q)^{1/q}}$ in the Lemma,
  $1 \geq \frac{E|X||Y|}{(E|X|^p)^{1/p}(E|Y|^q)^{1/q}}.$

Inequalities

- Theorem 4.7.3 (Cauchy-Schwartz inequality)
  For any two random variables X and Y,
  $|EXY| \le E|XY| \le (E|X|^2)^{1/2}(E|Y|^2)^{1/2}$
- (Liapounov's inequality)
  For $s > r$,
  $\{E|X|^r\}^{1/r} \le \{E(|X|^s)\}^{1/s}$, $1 < r < s < \infty$.
- (Minkowski's inequality) Let X and Y be any two random variables.
  Then for $1 \le p < \infty$,
  $[E|X+Y|^p]^{1/p} \le (E|X|^p)^{1/p} + (E|Y|^p)^{1/p}$
  $\langle Proof \rangle$

$$\begin{aligned} E|X+Y|^p &\le E|X||X+Y|^{p-1} + E|Y||X+Y|^{p-1} \\ &\le ((E|X|^p)^{1/p} + (E|Y|^p)^{1/p})(E|X+Y|^{(p-1)q})^{1/q} \\ &= ((E|X|^p)^{1/p} + (E|Y|^p)^{1/p})(E|X+Y|^p)^{1/q} \end{aligned}$$

Inequalities

- Definition
  $g(x)$: convex if $g(\lambda x + (1-\lambda)y) \leq \lambda g(x) + (1-\lambda)g(y)$ for $0 < \lambda < 1$.
- Theorem 4.7.7 (Jensen's Inequality)
  For any random variable X, if $g(x)$ is a convex function then $g(E(X)) \leq Eg(X)$.