

Exponential families

X is distributed according to an exponential family distribution if

$$p_\eta(x) = \exp\left(\sum_{i=1}^J \eta_i T_i(x) - A(\eta)\right) h(x) \quad (1)$$

is the density function of X (w.r.t. some measure μ).

Theorem

If X has distribution as in (1), then the moment-generating function of $T(X) = (T_1(X), \dots, T_J(X))$

$$M_T(u) = \frac{e^{A(\eta+u)}}{e^{A(\eta)}}$$

Proof.

By def

$$M_T(u) = E(e^{u_1 T_1 + \dots + u_J T_J})$$

~~$$= \int g(x) p_\eta(x) d\mu(x)$$~~

$$= \left(E(g(X)) = \int g(x) p_\eta(x) d\mu(x) \right)$$

$$= \int e^{\sum u_i T_i(x)} p_\eta(x) d\mu(x)$$

(1)

$$\begin{aligned}
&= \int e^{\sum u_i T_i(x)} e^{\sum v_i T_i(x) - A(y)} h(x) d\mu(x) \\
&= \int e^{\sum (u_i + v_i) T_i(x) - A(y)} h(x) d\mu(x) \\
&= e^{A(y+u) - A(y)} \int e^{\sum (u_i + v_i) T_i(x) - A(y+u)} h(x) d\mu(x) \\
&= e^{A(y+u) - A(y)} \cdot 1.
\end{aligned}$$

Ex. (Binomial moments)

Suppose $\mathbb{X} \in \text{Bin}(n, p)$.

$$\begin{aligned}
p_p(x) &= P(\mathbb{X} = x) \\
&= \binom{n}{x} p^x (1-p)^{n-x} \quad ; \quad x = 0, 1, \dots, n.
\end{aligned}$$

Rewrite to the canonical form expression

$$\begin{aligned}
&= \binom{n}{x} e^{x \log p + (n-x) \log(1-p)} \\
&= \binom{n}{x} e^{x \log(p/(1-p)) + n \log(1-p)}
\end{aligned}$$

This is an exponential family with

$$u_i = \log \frac{p}{1-p}, \quad T(x) = x$$

$$e^y = \frac{p}{1-p}$$

(2)

$$(1-p)e^{\eta} = p \quad ; \quad (e^{\eta} + 1)p = e^{\eta} ;$$

$$p = \frac{e^{\eta}}{1+e^{\eta}} \quad ; \quad \underline{1-p = \frac{1}{1+e^{\eta}}}$$

so that

$$\begin{aligned} A(\eta) &= -n \log(1-p(\eta)) \\ &= -n \log\left(\frac{1}{1+e^{\eta}}\right) \\ &= \underline{n \log(1+e^{\eta})} \end{aligned}$$

Therefore, the moment generating function is

$$\begin{aligned} M_T(u) \quad (&= M_X(u)) \\ &= e^{A(\eta+u) - A(\eta)} \\ &= e^{u \log(1+e^{\eta+u}) - n \log(1+e^{\eta})} \\ &= e^{u \log\left(\frac{1+e^{\eta+u}}{1+e^{\eta}}\right)} \\ &= \left(\frac{1+e^{\eta+u}}{1+e^{\eta}}\right)^n \end{aligned}$$

$$\left(e^{\eta} = \frac{p}{q} \right)$$

$$= \left(\frac{1+e^{\eta} \frac{p}{q}}{1+\frac{p}{q}} \right)^n = \left(\frac{q+e^{\eta} p}{p+q} \right)^n =$$

(3)

$$= \underline{(q + e^u p)^n}$$

Derive from this $E(\bar{X}), \text{Var}(\bar{X})$.

#1

Ex: (Normal moments)

$\bar{X} \in N(\xi, \sigma^2)$. Then the density, wrt length measure, is

$$\begin{aligned} p(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\xi)^2/2\sigma^2} && ; -\infty < x < \infty. \\ &= e^{-x^2/2\sigma^2 + x\xi/\sigma^2 - \xi^2/2\sigma^2} \cancel{+ \frac{1}{2}\log(2\pi\sigma^2)}^{1/2} \\ &\quad - \frac{1}{2}\log(2\pi\sigma^2) \end{aligned}$$

Thus, suppose that σ^2 fixed, then

$$\eta = \frac{\xi}{\sigma^2}$$

$$A(\eta) = \frac{\xi^2}{2\sigma^2} + \text{const} = \eta^2 \frac{\sigma^2}{2} + \text{const.}$$

Therefore, the moment generating function is

$$M(u) = M_{\bar{X}}(u) = \exp(A(\eta+u) - A(\eta))$$

~~$$= \exp((\eta+u)^2 \frac{\sigma^2}{2})$$~~

$$= \exp((\eta+u)^2 \frac{\sigma^2}{2} - \eta^2 \frac{\sigma^2}{2})$$

(4)

$$= \exp(\eta u^2 + \frac{u^2 v^2}{2})$$

$$= \exp(\eta u + \frac{u^2 v^2}{2})$$

Use this to obtain $E(X), \text{Var}(X)$.
#

Theorem

X is distributed according to exponential family distribution, then $T(X) = (T_1(X), \dots, T_J(X))$ is distributed according to exponential family with density

$$p_y(t) = \exp(\sum y_i t_i - A(y)) h(t)$$

with the appropriate measure.

Proof In testing statistical hypothesis, possibly do later.

#

Note that if X, Y independent, s -dimensional

$$\exp(\sum y_i T_i(x) - A(y)) h(x)$$

$$\exp(\sum y_j U_j(y) - C(y)) k(y)$$

their joint density is

$$\exp(\sum y_i (T_i(x) + U_i(y)) - (A(y) + C(y))) \cdot h(x) k(y).$$

This implies that

$$(T_1(X) + U_1(Y), \dots, T_s(X) + U_s(Y))$$

is s -dim exponential family distributed

$$\exp(\sum y_i v_i - (A(y) + C(y))) \cdot b(v)$$

If X_1, \dots, X_n are ~~i.i.d.~~^{indep.} 1-dim exponential with densities

$$\exp(y T_i(x_i) - A_i(y)) h_i(x_i)$$

then the joint density of X_1, \dots, X_n

$$\exp\left(y \sum_{i=1}^n T_i(x_i) - \sum_{i=1}^n A_i(y)\right) \tilde{h}(x)$$

then, again applying the previous theorem

we have that $\sum_{i=1}^n T_i(\bar{x}_i)$ is T -dir

exp-family distributed with density

$$\exp\left(yt - \sum_{i=1}^n A_i(y)\right) h(t)$$

sufficient statistics

Suppose that $\mathcal{F} = \{F_\theta : \theta \in \mathbb{R}\}$ is a family of d.f.'s.

Def. A statistic $T = T(\bar{x})$ is sufficient if

$$\bar{x} | T=t$$

does not depend on the parameter θ .

#

interpretation a) We don't know θ .

If T sufficient and if we know the distribution of $\bar{x} | T=t$, let \bar{x}' be a r.v. with that distribution, \bar{x}' distribution depends on t ($T = T(\bar{x}) = t$)

(7)

$$\underline{P_\theta(\bar{x}' \in A)} = E(1_{\{\bar{x}' \in A\}})$$

$$(P(B) = E(1_B))$$

$$= E[E(1_{\{\bar{x}' \in A\}} | T)]$$

$$(E(Y) = E(E(Y|X)))$$

$$= E[E(1_{\{X \in A\}} | T)]$$

$$= E(1_{\{X \in A\}}) = \underline{P_\theta(X \in A)}$$

b) since the distribution of $X|T=t$ does not depend on θ , and we assume it is known (we can with the use of clever probability theory derive it), we can use a random number generator to get an observation of \bar{x}' . This is possible even when we do not know the value of θ .

Recall we want to construct estimators
 $\delta(x)$

of θ . If x' is sample constructed
with help of random mechanism,
we can construct an estimator $\delta(x')$,
will be an estimator (does not depend
on θ), and it will be random.

Def. A randomized estimator of $f(\theta)$
is a rule which assigns to each outcome
 x of Σ a random variable $Y(x)$ with
known distribution.

#

To obtain a fixed number for the
estimator take an observation y of
 $Y(x)$, to get an estimate of
 $f(\theta)$.

⑨

Ex (Poisson).

Suppose \bar{X}_1, \bar{X}_2 ind. Poisson distributed with mean λ .

$$P(\bar{X}_1 = x_1, \bar{X}_2 = x_2)$$

~~$\lambda^{x_1} e^{-\lambda} / x_1!$ $\lambda^{x_2} e^{-\lambda} / x_2!$~~

$$= e^{-\lambda} \frac{\lambda^{x_1}}{x_1!} e^{-\lambda} \frac{\lambda^{x_2}}{x_2!}$$

$$= \frac{\lambda^{x_1+x_2}}{x_1! \cdot x_2!} e^{-2\lambda} \quad (1)$$

Let us look at $T(\bar{X}_1, \bar{X}_2) = \bar{X}_1 + \bar{X}_2$

(1) implies that

$$P(\bar{X}_1 = x_1 \mid T = t) = P(\bar{X}_1 = x_1 \mid \bar{X}_1 + \bar{X}_2 = t)$$

$$= \frac{P(\bar{X}_1 = x_1, \bar{X}_1 + \bar{X}_2 = t)}{P(\bar{X}_1 + \bar{X}_2 = t)}$$

$$= \frac{P(\bar{X}_1 = x_1, \bar{X}_2 = t - x_1)}{P(\bar{X}_1 + \bar{X}_2 = t)}$$

$$\begin{aligned}
 &= \frac{\lambda^{x_1 + (t-x_1)} e^{-2\lambda}}{\sum_{y=0}^t P(\bar{X}_1 = y, \bar{X}_2 = t-y)} / x_1! (t-x_1)! \\
 &= \frac{\lambda^{x_1 + (t-x_1)} e^{-2\lambda}}{\sum_{y=0}^t \lambda^y (t-y)!} / x_1! (t-x_1)! \\
 &= \frac{1}{x_1! (t-x_1)!} \cdot \left(\frac{1}{\sum_{y=0}^t \frac{1}{y! (t-y)!}} \right) \quad (\ast)
 \end{aligned}$$

We see that this does not depend on λ , therefore $T(\bar{X}_1, \bar{X}_2) = \bar{X}_1 + \bar{X}_2$ is a sufficient statistic.

How do we reconstruct (\bar{X}_1, \bar{X}_2) from T ?

Note that

$$\sum_{y=0}^t \frac{1}{y! (t-y)!} = \sum_{y=0}^t \binom{t}{y} \frac{1}{t!}$$

$$= \frac{1}{t!} \sum_{y=0}^t \binom{t}{y} = \frac{1}{t!} (1+1)^t$$

(11)

$$= \frac{1}{t!} 2^t$$

and

$$\frac{1}{x_1!(t-x_1)!} = \binom{t}{x_1} \frac{1}{t!}$$

\Rightarrow

$$(*) = \binom{t}{x_1} \frac{1}{2^t} = \binom{t}{x_1} \left(\frac{1}{2}\right)^{x_1} \left(\frac{1}{2}\right)^{t-x_1}$$

for the distribution of Σ_1 given
 $T = t$ is $\text{Bin}(t, 1/2)$.

to be continued . . .