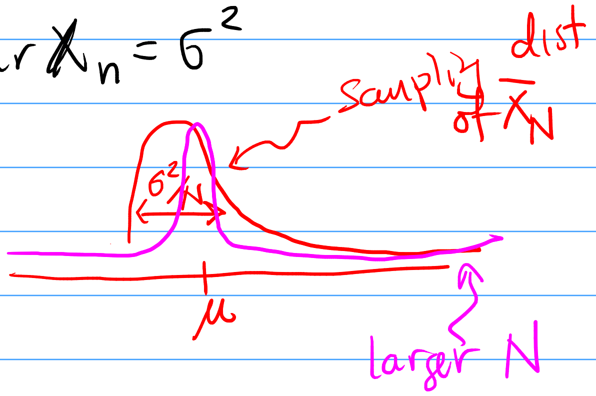
Lecture 2: Normal Stats and Exp Fams

Theorem: $X_n \overset{iid}{\sim} f$ so that

$$\mathbb{E}X_n = \mu \quad \text{and} \quad \operatorname{Var} X_n = \sigma^2$$

then
(1) $\mathbb{E}\bar{X}_N = \mu$
(2) $\operatorname{Var}(\bar{X}_N) = \sigma^2/N$
(3) $\mathbb{E}[S^2_{N-1}] = \sigma^2$



Pf. (1) $\mathbb{E}\bar{X}_N = \mathbb{E}\frac{1}{N}\sum_n X_n = \frac{1}{N}\sum_{n=1}^{N} \underbrace{\mathbb{E}[X_N]}_{\mu} = \frac{1}{N}\sum_{n=1}^{N}\mu = \frac{1}{N}N\mu$

$= \mu$

(2) $\operatorname{Var}(\bar{X}_N) = \operatorname{Var}\left(\frac{1}{N}\sum_{n=1}^{N} X_n\right)$

$= \frac{1}{N^2}\operatorname{Var}\left(\sum_n X_n\right)$

$= \frac{1}{N^2}\sum_n \underbrace{\operatorname{Var}(X_n)}_{\sigma^2}$

$= \frac{1}{N^2}\sum_{n=1}^{N}\sigma^2 = \frac{1}{N^2}N\sigma^2 = \sigma^2/N$

(3) $\mathbb{E}[S^2_{N-1}] = \mathbb{E}\left[\frac{1}{N-1}\sum_{n=1}^{N}(X_n - \bar{X}_N)^2\right]$

$= \frac{1}{N-1}\mathbb{E}\left[\sum_n (X_n - \bar{X})^2\right]$

$= \frac{1}{N-1}\mathbb{E}\left[\sum_n X_n^2 - N\bar{X}^2\right]$

recall: $\operatorname{Var} Z = \mathbb{E}(Z^2) - (\mathbb{E}Z)^2$

$\sum_n (X_n - \bar{X})^2 = \sum_n X_n^2 - N(\bar{X})^2$

$\mathbb{E}Z^2 = \operatorname{Var}(Z) + (\mathbb{E}Z)^2$

$$= \frac{1}{N-1}\left(\sum_n \mathbb{E}[X_n^2] - N\,\mathbb{E}[\bar{X}^2]\right)$$

$$\mathbb{E}[X_n^2] = Var(X_n) + \mathbb{E}[X_n]^2$$ (rearrange short-cut formula)

$$= \sigma^2 + \mu^2$$

$$\mathbb{E}[\bar{X}^2] = Var(\bar{X}) + (\mathbb{E}\bar{X})^2 =$$

$$= \sigma^2/N + \mu^2$$

$$= \frac{1}{N-1}\left(\sum_n(\sigma^2+\mu^2) - N(\sigma^2/N + \mu^2)\right)$$

$$= \frac{1}{N-1}\left(N\sigma^2 + \cancel{N\mu^2} - \sigma^2 - \cancel{N\mu^2}\right)$$

$$= \frac{1}{N-1}(N-1)\sigma^2 = \sigma^2$$

Theorem: If $X_n \overset{iid}{\sim} f$ then

$$M_{\bar{X}}(t) = \left(M(t/N)\right)^N$$

N ← sample size

M = mgf of any X_n individ.

Pf.

$$\bar{X} = \frac{1}{N}\sum_n X_n$$

Generic fact for MGF

$$M_{aZ+b}(t) = e^{tb}M_Z(at)$$

$$M_{\bar{X}}(t) = \mathbb{E}[e^{t\bar{X}}] = \mathbb{E}\left[e^{t\frac{1}{N}\sum_n X_n}\right]$$

$$e^a e^b = e^{a+b}$$

$$\prod_i e^{a_i} = e^{\sum_i a_i}$$

$$= \mathbb{E}\left[\prod_n e^{t/N\, X_n}\right]$$

$A \perp\!\!\!\perp B$ then $\mathbb{E}[AB] = (\mathbb{E}A)(\mathbb{E}B)$

$$= \prod_n \mathbb{E}\left[e^{t/N\, X_n}\right]$$

$= \prod_n M_{\cancel{X_n}}(t/N)$

$M_{X_n}(t/N) = \mathbb{E}\left[e^{\frac{t}{N}X_n}\right]$

$= M(t/N)^N$

Ex, $X_n \overset{iid}{\sim} \text{Gamma}(\alpha, \beta)$

$f(x) = \dfrac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \mathbb{1}(x>0)$

$M(t) = (1 - t/\beta)^{-\alpha}$

$M_{\bar{X}}(t) = M(t/N)^N$

$= \left[\left(1 - \frac{t/N}{\beta}\right)^{-\alpha}\right]^N$

$= \left(1 - \frac{t}{N\beta}\right)^{-N\alpha}$

repl. $\beta \rightsquigarrow N\beta$

$\alpha \rightsquigarrow N\alpha$

has form of MGF of Gamma$(N\alpha, N\beta)$

Ex, $\mathbb{E}[\bar{X}] = \frac{N\alpha}{N\beta} = \frac{\alpha}{\beta}$.

$\bar{X} \sim \text{Gamma}(N\alpha, N\beta)$.

Theorem: $\bar{X}$ and S^2 for normal

Let $X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

(prove) ① $\bar{X} \sim N(\mu, \sigma^2/N)$

(later) ② $\bar{X} \perp S^2$

(sketch) ③ $\frac{N-1}{\sigma^2} S_{N-1}^2 \sim \chi^2(N-1)$

$N-1 = \mathbb{E}\left[\frac{N-1}{\sigma^2} S^2\right] = \frac{N-1}{\sigma^2}\mathbb{E}[S^2]$

$\Rightarrow \mathbb{E}[S^2] = \sigma^2$

chi-squared dist w/ $N-1$ degrees of freedom.

Pf.

(1) $M_{\bar{X}}(t) = M(t/N)^N$

MGF of Normal is $M(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$

$$= \exp\left(\mu(t/N) + \frac{\sigma^2 (t/N)^2}{2}\right)^N$$

$$= \exp\left(\frac{\mu t}{N} + \frac{\sigma^2 t^2}{2N^2}\right)^N$$

$$= \exp\left(\frac{N\mu t}{N} + \frac{N\sigma^2 t^2}{2N^2}\right)$$

$$= \exp\left(\mu t + \frac{\sigma^2/N \, t^2}{2}\right)$$

$(e^a)^b = e^{ab}$

= MGF of $N(\mu, \sigma^2/N)$

So $\bar{X} \sim N(\mu, \sigma^2/N)$

Chi-Squared Dist $X \sim \chi^2(k)$

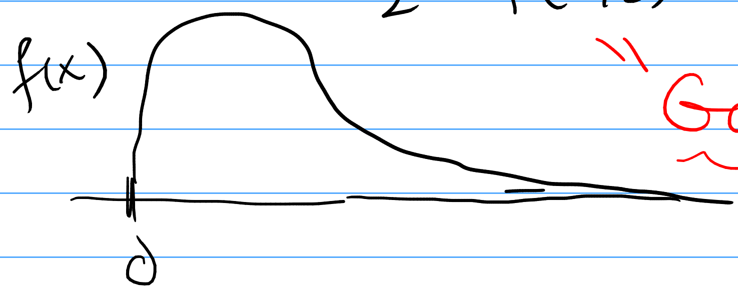
One parameter k – degrees of freedom

$E[X] = k$

$Var(X) = 2k$

$$f(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2} \mathbb{1}(x>0)$$

" Gamma($\alpha = k/2, \beta = 1/2$)



Facts: (1) $Z \sim N(0,1)$ then $Z^2 \sim \chi^2(1)$

(2) $Z_1 \sim N(0,1)$, $Z_2 \sim N(0,1)$, $Z_1 \perp\!\!\!\perp Z_2$ then $Z_1^2 + Z_2^2 \sim \chi^2(2)$

Generically: $Z_i \overset{iid}{\sim} N(0,1)$ then $\sum_{i=1}^{N} Z_i^2 \sim \chi^2(N)$

③ $Y_n \overset{\text{indep}}{\sim} \chi^2(k_n)$ then $\sum_n Y_n \sim \chi^2\left(\sum_n k_n\right)$

Sums of indep χ^2 is χ^2 where you sum the DOF

Sketch of fact ③

$X_n \overset{\text{iid}}{\sim} N(\mu,\sigma^2)$

$\bar{X} \sim N(\mu, \sigma^2/N)$

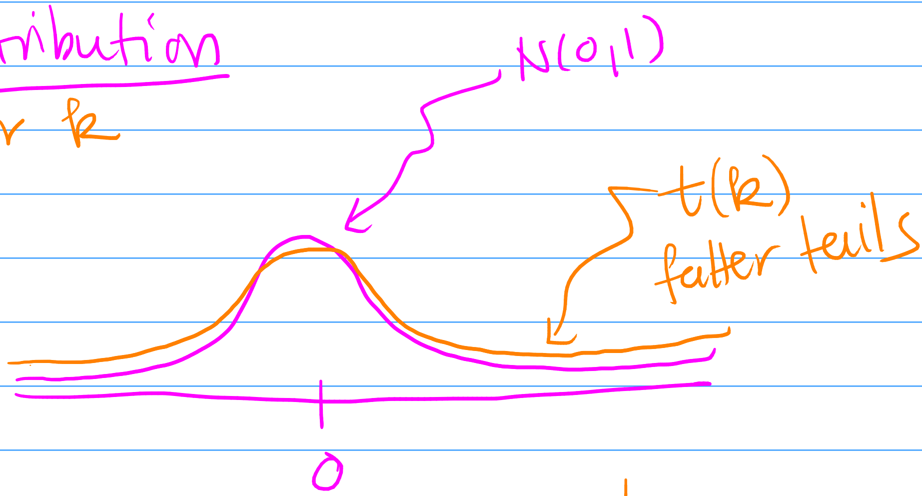
$$S^2_{N-1} = \frac{1}{N-1}\sum_n (\underbrace{X_n - \bar{X}}_{\text{almost like } Z_n \sim N(0,\sigma^2)})^2$$

$$\frac{N-1}{\sigma^2} S^2_{N-1} = \sum_{n=1}^{N} \Big(\underbrace{\frac{X_n - \bar{X}}{\sigma}}_{\text{almost } N(0,1)}\Big)^2 \approx \chi^2(N-1)$$

---

t-distribution

One parameter k = dof



$$f(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{k\pi}\,\Gamma(k/2)} \frac{1}{\left(1 + x^2/k\right)^{\frac{k+1}{2}}} \quad \text{for all } x \in \mathbb{R}$$

Fact: $\left\{\begin{array}{l} U \sim N(0,1) \\ V \sim \chi^2(k) \end{array}\right\} U \perp\!\!\!\perp V$

then $\boxed{\dfrac{U}{\sqrt{V/k}} \sim t(k)}$ ← can prove w/ a biv. transf. from 4.5)

If $X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ then

① $\bar{X} \sim N(\mu, \sigma^2/N)$ ③ $\frac{N-1}{\sigma^2} S^2 \sim \chi^2(N-1)$

and ② $\bar{X} \perp\!\!\!\perp S^2$

Intro stats:

$$t = \frac{\bar{X} - \mu}{S/\sqrt{N}} \sim t(N-1)$$

t-stat.

$$U = \frac{\bar{X} - \mu}{\sigma/\sqrt{N}} \sim N(0,1)$$

why? $\bar{X} \sim N(\mu, \sigma^2/N) \Rightarrow \bar{X} - \mu \sim N(0, \sigma^2/N)$

$\Rightarrow \dfrac{\bar{X} - \mu}{\sigma/\sqrt{N}} \sim N(0,1)$

Know:

$$V = \frac{N-1}{\sigma^2} S^2 \sim \chi^2(N-1)$$

if $\bar{X} \perp\!\!\!\perp S^2 \Rightarrow U \perp\!\!\!\perp V$

So $\dfrac{U}{\sqrt{V/_{N-1}}} \sim t(N-1)$

$$\frac{U}{\sqrt{V/_{N-1}}} = \frac{\dfrac{\bar{X}-\mu}{\sigma/\sqrt{N}}}{\sqrt{\frac{N-1}{\sigma^2}S^2/_{N-1}}} = \frac{\dfrac{\bar{X}-\mu}{\cancel{\sigma}/\sqrt{N}}}{S/\cancel{\sigma}} = \frac{\bar{X}-\mu}{S/\sqrt{N}} = t\text{-stat}$$

so as claimed it has a $t(N-1)$.

<u>Probability</u>: Given $X_n \overset{iid}{\sim} f_\theta$ ← parameter

if we know $\theta = 5$

Calculate $\mathbb{P}(X_n = \cdots)$

<u>Statistics</u>:

If I observe $X_n \overset{iid}{\sim} f_\theta$ but I don't know θ.

How can I estimate it?

What can I say about the estimator?

<u>Ex.</u> $X_n \overset{iid}{\sim} N(\mu, 1)$

How can I estimate μ? $\bar{X}$?

How good is $\bar{X}$ as an estimator of μ?

Ex. $X_n \overset{iid}{\sim} \text{Exp}(\lambda)$ when $\lambda > 0$ is unknown.

How can I estimate λ?

Generally, we'll work w/ parameterized families of dists:

e.g. $N(\mu, \sigma^2)$ where $\mu \in \mathbb{R}$, $\sigma^2 > 0$

$\text{Exp}(\lambda)$ when $\lambda > 0$

$U(0, \theta)$ when $\theta > 0$

Exponential Families

θ = parameter

space of possible params

Assume we have a family of dists parameterized by some param $\theta \in \Theta \subset \mathbb{R}$

one-dim'l params

so that

$$X_n \overset{iid}{\sim} f_\theta$$

and assume that

depend on θ but not $\underset{\sim}{x}$

$$f(\underset{\sim}{x}) = h(\underset{\sim}{x}) c(\theta) \exp\left(T(\underset{\sim}{x}) w(\theta)\right)$$

joint PMF/PDF

depend on $\underset{\sim}{x}$ not θ

then we say that the X_ns have an exp. family

Ex. $X_n \overset{iid}{\sim} \mathrm{Pois}(\underset{\theta}{\lambda})$ $\quad \underset{\sim}{\lambda > 0}$ $\ominus$

$$f(\underset{\sim}{x}) = \prod_{n=1}^{N} f(x_n) = \prod_n \underbrace{\frac{1}{x_n!}\lambda^{x_n} e^{-\lambda} \mathbb{1}(x_n \in \mathbb{N}_0)}_{\mathrm{Pois}(\lambda)}$$

$$= \left(\prod_n \frac{1}{x_n!}\right)\left(\prod_n \lambda^{x_n}\right)\left(\prod_n e^{-\lambda}\right)\prod_n \mathbb{1}(x_n \in \mathbb{N}_0)$$

$$= \left(\prod_n \frac{1}{x_n!}\right)\left(\prod_n \mathbb{1}(x_n \in \mathbb{N}_0)\right) \lambda^{\sum_n x_n} e^{-N\lambda}$$

$$f(\underset{\sim}{x}) = \underbrace{\left(\prod_n \frac{1}{x_n!}\right)\left(\prod_n \mathbb{1}(x_n \in \mathbb{N}_0)\right)}_{h(\underset{\sim}{x})} \exp\big(\underbrace{(\textstyle\sum_n x_n)}_{T(\underset{\sim}{x})}\underbrace{\log\lambda}_{w(\lambda)}\big) \underbrace{e^{-N\lambda}}_{c(\lambda)}$$

So X_n jointly have an exp. fam.