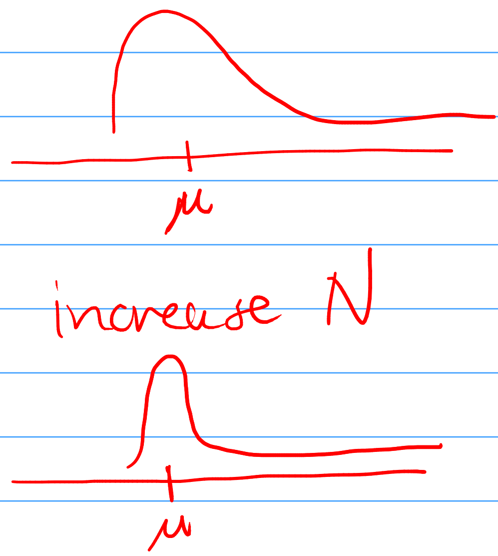
Lecture 2: Normal Stats and Exponential Families

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

$$\mu = EX_n \quad , \quad \sigma^2 = Var(X_n)$$

then

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



Pf. ① $\mathbb{E}[\bar{X}] = \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N} X_n\right] = \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}[X_n]$

$$= \frac{1}{N}\sum_{n=1}^{N}\mu = \frac{1}{N}N\mu = \mu$$

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

by independence

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

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

recall: $Var(z) = \mathbb{E}[z^2] - \mathbb{E}[z]^2$

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

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

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

$Var(z) = \mathbb{E}[z^2] - \mathbb{E}[z]^2$

rearrange: $\mathbb{E}[z^2] = Var(z) + \mathbb{E}[z]^2$

$\mathbb{E}[X_n^2] = Var(X_n) + \mathbb{E}[X_n]^2 = \sigma^2 + \mu^2$

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

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

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

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

<u>Theorem</u>: If $X_n \overset{iid}{\sim} f$

MGF of any X_n

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

$$M_X(t) = \mathbb{E}[e^{tX}]$$

pf. $\bar{X} = \frac{1}{N}\sum_{n=1}^{N} X_n$

Facts! - $M_{aX+b}(t) = M_X(at)e^{tb}$

$*\, A \perp\!\!\!\perp B$ then $E[AB] = E[A]E[B]$

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

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

$e^{\sum a_n} = \prod_n e^{a_n}$

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

$\mathbb{E}[\prod A_n] = \prod_n \mathbb{E}[A_n]$

$$= \prod_n \underbrace{\mathbb{E}[e^{t/N\, X_n}]} \quad \text{(by independence)}$$

$M(t) = \mathbb{E}[e^{tX_n}]$

$M(t/N) = \mathbb{E}[e^{t/N\, X_n}]$

$$= \prod_n M(t/N)$$

$$= M(t/N)^N$$

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

What is the dist of $\bar{X}$?

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

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

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

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

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

looks like MGF of Gama

repl. α w/ $N\alpha$

β w/ $N\beta$

So MGF of $\bar{X}$ is MGF of Gamma$(\alpha N, \beta N)$

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

Theorem: $\bar{X}$ and S^2 for Normal Dists.

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

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

(later) ② $\bar{X} \perp\!\!\!\perp S^2$

(sketch) ③ $\frac{N-1}{\sigma^2} S^2 \sim \chi^2(N-1)$ chi-squared dist w/ $N-1$ degrees of freedom

---

Pf. ① $M_{\bar{X}}(t) = M(t/N)^N$ | $M(t) = \exp(\mu t + \sigma^2 t^2/2)$

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

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

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

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

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

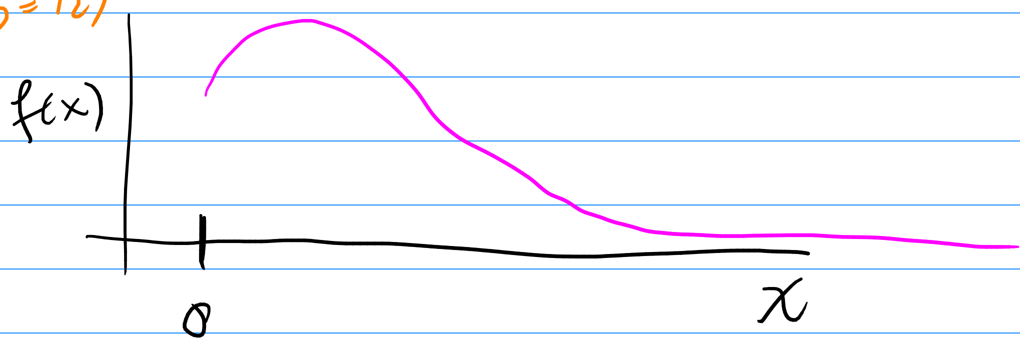
So $\bar{X} \sim N(\mu, \sigma^2/N)$

Chi-Squared Distributions

$X \sim \chi^2(k)$ DOF, $k>0$

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

$= \text{Gamma}(\alpha = k/2, \beta = 1/2)$



$\mathbb{E}[X] = k$, $\text{Var}(X) = 2k$

Facts!

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

2. If $Z_1 \sim N(0,1)$, $Z_2 \sim N(0,1)$, $Z_1 \perp\!\!\!\perp Z_2$ then

$$Z_1^2 + Z_2^2 \sim \chi^2(2)$$

Generally if $Z_n \overset{iid}{\sim} N(0,1)$ then

$$\sum_{n=1}^{N} Z_n^2 \sim \chi^2(N)$$

3. If $Y_n \overset{indep}{\sim} \chi^2(k_n)$ then

$$\sum_{n=1}^{N} Y_n \sim \chi^2\left(\sum_{n=1}^{N} k_n\right)$$

Argument for 3 in above theorem

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

$$S^2 = \frac{1}{N-1} \sum_{n=1}^{N} (X_n - \bar{X})^2 \quad \text{so} \quad \frac{N-1}{\sigma^2} S^2 = \frac{1}{\sigma^2} \sum_n (X_n - \bar{X})^2$$

$$= \sum_n \left(\frac{X_n - \bar{X}}{\sigma} \right)^2$$

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

$\rightarrow X_n - \mu \sim N(0, \sigma^2)$

$\rightarrow \frac{X_n - \mu}{\sigma} \sim N(0,1)$

$$\approx \sum_n \left(\frac{X_n - \mu}{\sigma} \right)^2$$

$$= \sum_n N(0,1)^2 \sim \chi^2(N)$$

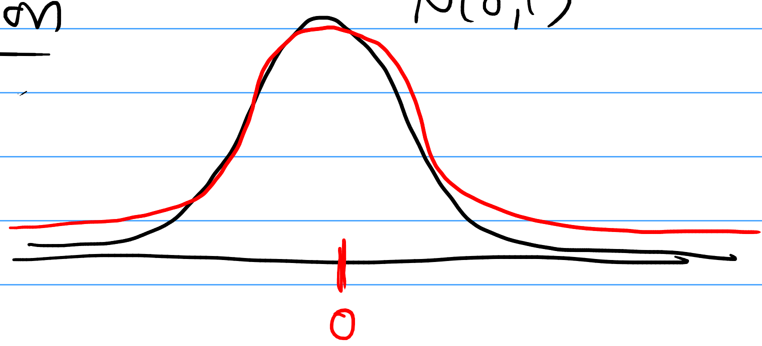
can't exactly do this

Penalty for using $\bar{X}$ not μ is that DOF reduced from N to $N-1$

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

---

t-distribution



has one parameter k = DOF

$$f(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{k\pi}} \frac{1}{\Gamma(k/2)} \frac{1}{\left(1 + x^2/k\right)^{\frac{k+1}{2}}}$$

$\forall x \in \mathbb{R}$

Fact: $U \sim N(0,1)$, $V \sim \chi^2(k)$ $\Big\}$ $U \perp\!\!\!\perp V$

then $\dfrac{U}{\sqrt{V/k}} \sim t(k)$

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

① $\boxed{\bar{X} \sim N(\mu, \sigma^2/N)}$ ② $\bar{X} \perp\!\!\!\perp S^2$

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

So $\dfrac{\bar{X} - \mu}{S/\sqrt{N}} \sim t(N-1)$ here $S = \sqrt{S^2}$

Why?

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

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

$$\frac{U}{\sqrt{V/_{N-1}}} = \frac{\frac{\bar{X}-\mu}{\cancel{\sigma}/\sqrt{N}}}{\sqrt{\frac{\cancel{N-1}}{\cancel{\sigma^2}} S^2 / \cancel{N-1}}} = \frac{\bar{X}-\mu}{\sqrt{S^2}/\sqrt{N}} = \frac{\bar{X}-\mu}{S/\sqrt{N}} \sim t(N-1)$$

$t(N-1)$