# Chi-Squared as Conjugate Prior

- Consider $x_i \mid \lambda \sim \text{Poisson}(\lambda)$ i.i.d

  In other words $P(x_i = k \mid \lambda) = \dfrac{\lambda^k e^{-\lambda}}{k!}$

- Assume the following prior on $\lambda$.

$$m\lambda \sim \chi^2_k$$

what does that mean?

$$f(m\lambda) = \frac{1}{2^k \Gamma(\frac{k}{2})} (m\lambda)^{\frac{k}{2}-1} e^{-m\lambda/2}$$

So $f(\lambda) = \dfrac{m}{2^k \Gamma(\frac{k}{2})} (m\lambda)^{k/2-1} e^{-m\lambda/2}$ [Change of variable]

$$\frac{f(m\lambda)}{\left|\frac{d\lambda}{d(m\lambda)}\right|} = f(m\lambda)\left|\frac{d(m\lambda)}{d\lambda}\right| = m f(m\lambda)$$

- Find $f(\lambda \mid x_1 \ldots x_n)$

$$f(\lambda \mid x_1 \dots x_n) \propto P(x_1 \dots x_n \mid \lambda)\, f(\lambda)$$

$$\propto \lambda^{x_1} e^{-\lambda} \dots \lambda^{x_n} e^{-\lambda} (m\lambda)^{\frac{k}{2}-1} e^{-m\lambda/2}$$

$$= \lambda^{T} e^{-n\lambda} (m\lambda)^{\frac{k}{2}-1} e^{-m\lambda/2}$$

where $T = \sum_{i=1}^{n} x_i$

$$f(\lambda \mid x_1 \dots x_n) \propto \left[(m+2n)\lambda\right]^{\frac{k+2T}{2}-1} e^{-(m+2n)\frac{\lambda}{2}}$$

so $(m+2n)\lambda \mid x_1 \dots x_n \sim \chi^2_{k+2T}$

# The Normal case

- Suppose $x_i \mid \sigma^2 \sim N(\mu, \sigma^2)$ are i.i.d (Conditioned on $\sigma^2$) and $\mu$ is known.
- We can show that $S_0/\sigma^2 \sim \chi^2_k$ is a Conjugate prior.

First, observe that $f(\sigma^2 \mid x_1 \dots x_n) \propto f(x_1 \dots x_n \mid \sigma^2) f(\sigma^2)$

$$\propto \frac{1}{\sigma^n} e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}} f(\sigma^2)$$

$$= \left(\frac{1}{\sigma^2}\right)^{n/2} e^{-\frac{\sum(x_i-\mu)^2}{2\sigma^2}} f(\sigma^2)$$

Then, since everything seems to be in terms of $\frac{1}{\sigma^2}$, then

$$f\left(\frac{1}{\sigma^2} \,\Big|\, x_1 \dots x_n\right) \propto \left(\frac{1}{\sigma^2}\right)^{n/2} e^{-\frac{\sum(x_i-\mu)^2}{2\sigma^2}} f\left(\frac{1}{\sigma^2}\right)$$

Finally, if $S_0/\sigma^2 \sim \chi^2_k$, then $f\left(\frac{1}{\sigma^2}\right) \propto \left(\frac{1}{\sigma^2}\right)^{\frac{k}{2}-1} e^{-\frac{S_0}{2\sigma^2}}$

$$f\left(\frac{1}{\sigma^2} \,\middle|\, x_1 \dots x_n\right) \propto \left(\frac{1}{\sigma^2}\right)^{\frac{k+n}{2}-1} e^{-\frac{(S+S_0)}{2\sigma^2}}$$

which means $(S+S_0)/\sigma^2 \sim \chi^2_{k+n}$

where $S = \sum_{i=1}^{n} (x_i - \mu)^2$

## Interpretation of $k$ and $S_0$ :

- Since $k$ is added to $n$, $k$ can be interpreted as the number of "observations" that lead to the prior
- $S_0$ is the sum $\sum (x_i - \mu)^2$ where $x_i$'s are the $k$ observations.

## Example:

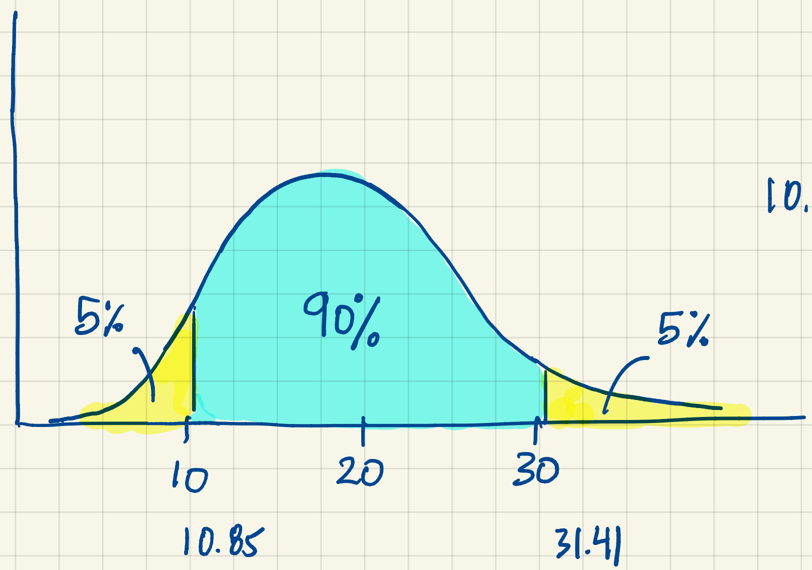
| | | | | |
|---|---|---|---|---|
| 9 | 18 | 21 | 26 | 14 |
| 18 | 22 | 27 | 15 | 19 |
| 22 | 29 | 15 | 19 | 24 |
| 30 | 16 | 20 | 24 | 32 |

$\mu = 22$

$n = 20 \qquad S = 664$

Without any prior knowledge $\kappa = S_0 = 0$

$$\frac{(S+S_0)}{\sigma^2} \Big| x_1 \cdots x_n \sim \chi^2_{n+\kappa} \Rightarrow \frac{664}{\sigma^2} \Big| x_1 \cdots x_{20} \sim \chi^2_{20}$$




$$10.85 < \frac{644}{\sigma^2} < 31.41$$

Approximately

$20 \leq \sigma^2 \leq 60$

with 90% Prob.

What kind of prior is $\frac{S_0}{\sigma^2} \sim \chi^2_k$ when $S_0 = k = 0$ ?

$$S_0/\sigma^2 \sim \chi^2_k \Rightarrow f\left(\frac{1}{\sigma^2}\right) \propto \left(\frac{1}{\sigma^2}\right)^{k/2 - 1} e^{-S_0/2\sigma^2}$$

so $f\left(\frac{1}{\sigma^2}\right) \propto \sigma^2$

This has the form $f(x) = \frac{a}{x}, \; x \geq 0$

improper prior.

$$\int_0^\infty \frac{a}{x}\,dx = a \ln x \Big|_0^\infty = \infty + \infty = \infty$$

And yet, this prior resulted in a valid posterior.

# Connection to classical approach.

Consider the random variable $\frac{x_i - \mu}{\sigma} \mid \sigma^2 \sim N(0,1)$

Then $\frac{(x_i-\mu)^2}{\sigma^2} \mid \sigma^2 \sim \chi^2_1$

If $x_1 \dots x_n$ are independent conditioned on $\sigma^2$, then

$$\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{\sigma^2} \sim \chi^2_{20} \text{, so } \quad S/\sigma^2 \sim \chi^2_n$$

We got what we knew already!

## What if both $\mu$ and $\sigma^2$ are unknown?

$x_i \mid \mu, \sigma^2 \sim N(\mu, \sigma^2)$ are i.i.d conditioned on $\mu$ and $\sigma^2$

Define: $\bar{x} = \sum_{i=1}^{n} x_i / n$ (average) $E[\bar{x}] = \mu$

$s^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 / (n-1)$ is a "good" estimate of $\sigma^2$

because we can show

$E[s^2] = \sigma^2$

Let's reconsider the z-test

Recall $\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

But, the term $\frac{\bar{x} - \mu}{s/\sqrt{n}}$ cannot be claimed to be $\sim N(0,1)$

- while $\mu = E[x]$

- $s^2 \neq E[(x-\mu)^2]$

# The t-distribution (student)

- $Z \sim N(0,1)$
- $V \sim \chi^2_K$
- $Z$ and $V$ are independent

Then $t = \dfrac{Z}{\sqrt{V/K}}$ satisfies $f(t) \propto \left(1+\dfrac{t^2}{K}\right)^{-\frac{K+1}{2}}$ [look up constant]

Observation: when $K$ is large, this is almost Normal.

$$\lim_{K\to\infty} \left(1+ \underbrace{\frac{K+1}{K}}_{\downarrow\, 1} \cdot \frac{t^2}{K+1}\right)^{-\frac{K+1}{2}} = e^{-t^2/2}$$

$$\lim_{n\to\infty} \left(1+\frac{x}{n}\right)^n \to e^x$$

$E[t] = 0$ if $K > 1$

$$\mathrm{Var}[t] = \begin{cases} \dfrac{K}{K-2} & K > 2 \\ \infty & 1 < K \leq 2 \end{cases}$$

Fisher proved the following for Normal samples

- $(n-1)s^2/\sigma^2 \sim \chi^2_{n-1}$

- $\bar{x}$ and $s^2$ are independent (only Normal satisfies this)

- (we know) $\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

So $$\frac{\bar{x}-\mu}{s/\sqrt{n}} \sim t_{n-1} \quad (t\text{-distributed with degree } n-1)$$

Remark: The text in Note 8 contains proofs for the first two facts when $n=2$.

# Two teams revisited (unknown $\sigma^2$)

$$\frac{(n-1)s_x^2}{\sigma_x^2} \sim \chi^2_{n-1} \qquad \frac{(m-1)s_y^2}{\sigma_y^2} \sim \chi^2_{m-1}$$

So $$\frac{(n-1)s_x^2}{\sigma_x^2} + \frac{(m-1)s_y^2}{\sigma_y^2} \sim \chi^2_{n+m-2} \quad [\text{sum of ind. } \chi^2]$$

From before $$\frac{\bar{x} - \bar{y} - (\mu_x - \mu_y)}{\sqrt{\sigma_x^2/n + \sigma_y^2/m}} \sim N(0,1)$$

Assuming $\sigma_x^2 = \sigma_y^2$ [Though both are unknown], testing $\mu_x = \mu_y$ we get

$$\frac{\bar{x} - \bar{y}}{\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right)\frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}}} \sim t_{n+m-2}$$

[Note 8 has an example]