Last time I claimed (min, max) sufficient for Uniform$(0,\vartheta)$. It is enough with max:

$$L(x|\vartheta) = \frac{1}{\vartheta^n} \cdot \mathbb{1}_{(-\infty,\vartheta)}(\max) \cdot \mathbb{1}_{(0,\infty)}(\min)$$

so $h(x) = \mathbb{1}_{(0,\infty)}(\min)$ fulfills factorization theorem!

Conditional probability

Recall that the conditional density is defined as

$$f(x|y) = f(x,y)/f(y) \qquad \mathbb{P}(x|y) = \mathbb{P}(x,y)/\mathbb{P}(y).$$

We define

$$\mathbb{E}\left[g(X) \mid Y=y\right] = \int_{-\infty}^{\infty} g(x)\, f(x|y)\, dx$$

or

$$\mathbb{E}\left[g(X) \mid Y=y\right] = \sum_x g(x)\, \mathbb{P}(x|y)$$

We can define a conditional variance:

$$\mathbb{V}[X \mid Y=y] = \mathbb{E}[X^2|y] - \mathbb{E}[X|y]^2$$

These conditional expectations and variances have properties important for our next major result.

Theorem $\mathbb{E}X = \mathbb{E}_Y\left[\mathbb{E}[X|Y]\right]$

Proof $\mathbb{E}_Y\left[\mathbb{E}[X|Y]\right] = \int_{-\infty}^{\infty} \mathbb{E}[X|Y=y]\, f(y)\, dy =$

$$= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} x\, f(x|y)\, dx \right) f(y)\, dy =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f(x|y)\, f(y)\, dx\, dy =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f(x,y)\, dx\, dy =$$

$$= \int_{-\infty}^{\infty} x\, f(x)\, dx = \mathbb{E}X. \qquad \blacksquare$$

Theorem $\mathbb{V}X = \mathbb{E}_Y\left[\mathbb{V}[X|Y]\right] + \mathbb{V}_Y\left[\mathbb{E}[X|Y]\right]$

Proof $\mathbb{E}[\mathbb{V}(X|Y)] + \mathbb{V}[\mathbb{E}(X|Y)] =$

$$= \mathbb{E}\left[\mathbb{E}[X^2|Y]\right] \underbrace{- \mathbb{E}\left[\mathbb{E}[X|Y]^2\right] + \mathbb{E}\left[\mathbb{E}[X|Y]^2\right]}_{=0} - \mathbb{E}\left[\mathbb{E}(X|Y)\right]^2$$

$$= \mathbb{E}\left[\mathbb{E}[X^2|Y]\right] - \mathbb{E}\left[\mathbb{E}[X|Y]\right]^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{V}X. \qquad \blacksquare$$

Our main theorem today helps us find best possible unbiased estimators.

Theorem (Rao-Blackwell)
Let $\hat{\vartheta}$ be unbiased for ϑ w finite variance.
If U is a sufficient statistic for ϑ, define $\vartheta^* = \mathbb{E}[\hat{\vartheta} \mid U]$.
Then, for all ϑ:

$$\mathbb{E}[\vartheta^*] = \vartheta \text{ and } \mathbb{V}(\vartheta^*) \leq \mathbb{V}(\hat{\vartheta}).$$

Proof First, we observe that by sufficiency, ϑ^* does not depend on ϑ, so ϑ^* is a statistic.

Next, observe $\mathbb{E}\vartheta^* = \mathbb{E}_U[\mathbb{E}[\hat{\vartheta} \mid U]] = \mathbb{E}\hat{\vartheta} = \vartheta$.

Furthermore $\mathbb{V}[\hat{\vartheta}] = \mathbb{V}[\mathbb{E}[\hat{\vartheta} \mid U]] + \mathbb{E}[\mathbb{V}[\hat{\vartheta} \mid U]] =$

$$= \mathbb{V}\vartheta^* + \mathbb{E}[\mathbb{V}[\hat{\vartheta} \mid U]] \geq \mathbb{V}\vartheta^*. \qquad \blacksquare$$

We can apply Rao-Blackwell again to ϑ^*, but this is unlikely to improve things; say that our first result was some $h(U)$. Then the next application would produce

$$\mathbb{E}[h(U) \mid U] = h(U).$$

The way to improve is through a <u>new</u> statistic.

Minimal Sufficiency

If T is sufficient for ϑ and $T = f(T')$ for some other statistic T', then clearly T' is also sufficient.

An "essential" sufficient statistic would be one that can be generated from every possible sufficient statistic.

Def A sufficient statistic T is minimal if for every other sufficient statistic T', there is some f such that
$$T = f(T') \text{ (almost everywhere)}$$

Theorem Suppose T is sufficient for ϑ and for datasets x, y:
$$\mathcal{L}(x|\vartheta) \propto \mathcal{L}(y|\vartheta) \text{ implies } T(x) = T(y)$$
Then T is minimal sufficient

Proof sketch Suppose T' is sufficient. If there is no f s.t. $f(T') = T$, then there are datasets x, y s.t. $T'(x) = T'(y)$ and $T(x) \neq T(y)$.
But then
$$\mathcal{L}(x|\vartheta) = g'(\vartheta, T'(x))\, h'(x) \propto g'(\vartheta, T'(y))\, h'(y) = \mathcal{L}(y|\vartheta)$$
because $T'(x) = T'(y)$ it follows that they differ by a factor $h'(x)/h'(y)$.

By the assumption in the theorem, $T(x) = T(y)$ follows. $\blacksquare$

Lehmann-Scheffé

Pick two points x, y. If the likelihood ratio $\frac{L(x|\vartheta)}{L(y|\vartheta)}$ is independent of ϑ if and only if some $g(x) = g(y)$, then g is a minimal sufficient statistic for ϑ.

Example $Y_1, \dots, Y_n \sim \text{Bernoulli}(p)$ iid.

$$L(Y|p) = \left(\frac{p}{1-p}\right)^{\sum Y_i} \cdot (1-p)^n$$

so $$\frac{L(x|p)}{L(y|p)} = \left(\frac{p}{1-p}\right)^{\sum x_i - \sum y_i}.$$

p vanishes from this ratio exactly when $\sum x_i - \sum y_i = 0$.

Completeness

A statistic U is complete for ϑ if

$\mathbb{E}[f(U)] = 0$ implies $f(U) = 0$.

Theorem If T is complete and sufficient, then T is minimal sufficient.

Proof Let T' be minimal sufficient. Then there is a function $f(T) = T'$.
Define $g(T') = \mathbb{E}[T|T']$. $\mathbb{E}[g(T')] = \mathbb{E}T$ so
$\mathbb{E}[g(T') - T] = \mathbb{E}[g(f(T)) - T] = 0$.
By completeness, $g(T') - T = 0$ so $T = g(T')$ shows that T too is minimal sufficient. ∎

Def An exponential family is <u>full rank</u> if $T(x)$ does not have linear dependencies between entries.

Theorem The sufficient statistic T of a full rank exponential family is complete.

Def A statistic V whose distribution does not depend on ϑ is called <u>ancillary</u>.

Theorem (Basu) If T is complete and sufficient, and V is ancillary, then T and V are independent.

Example $\quad X_1, \dots, X_n$ iid $N(\mu, \sigma^2)$, σ^2 known.

$$L(x \mid \mu, \sigma^2) = \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}^{\,n}}}_{A} \exp\left[\underbrace{\frac{n\mu}{\sigma^2}}_{\eta} \underbrace{\bar{x}}_{T} - \underbrace{\frac{n\mu^2}{2\sigma^2}}_{A} \underbrace{- \frac{1}{2\sigma^2} \sum x_i^2}_{B} \right]$$

This is a full rank family, so $\bar{X}$ is complete sufficient.

Claim The sample variance is ancillary

Proof Let $Y_i = X_i - \mu$. Then $Y_i \sim N(0, \sigma^2)$.

$\bar{Y} = \bar{X} - \mu$ so $Y_i - \bar{Y} = X_i - \bar{X}$ and it follows that

$\frac{1}{n-1} \sum (X_i - \bar{X})^2 = \frac{1}{n-1} \sum (Y_i - \bar{Y})^2$ and the Y_i don't depend on μ.

Conclusion By Basu's theorem, $\bar{X}$ and S^2 are independant!

Exercise 9.68

If U is complete sufficient for ϑ.

Claim There is a unique Minimum Variance Unbiased Estimator:

Suppose $g_1(U)$ and $g_2(U)$ are both unbiased.
Then

$$\mathbb{E}[g_1(U)-g_2(U)] = \vartheta-\vartheta=0$$

so from completeness follows

$$g_1(U)-g_2(U)=0 \quad \text{ie} \quad g_1(U)=g_2(U). \text{ a.e.}$$