36-705 (07/07/2020 07/11/2020)

Youtube lecture 12/10/16

- Review lecture

- Test II preparation

- convergence
- delta method
- sufficiency

- point estimation - MOM, MLE, Bayes estimator
- MSE, consistency
- Risk function → MSE (i.e. L_2 loss).

* Not included ⟶ VC theory, minimax, score, Fisher information, probability inequalities (Test I).

- Prepare a cheat sheet; not a memory test

Q: How do you find minimal sufficient statistics?

LW: Ratio method. Find joint distribution on 2 different datasets; find what makes parameter drop out.

- Likelihood function forms a partition that creates an equivalence class

$(X_1, \ldots, X_n) \sim (Y_1, \ldots, Y_n) \qquad L_{x_1,\ldots,x_n}(\theta) \sim L_{y_1,\ldots,y_n}(\theta)$

3:05

- That partition

- Example: $X_1, \ldots, X_n \sim N(\theta, 1)$, $Y_1, \ldots, Y_n \sim N(\theta, 1)$

- Likelihoods are proportional iff $\bar{X}_n = \bar{Y}_n$
- Hence $\bar{X}_n$ is minimal sufficient

- Maximum likelihood estimator needs not always be sufficient
- Likelihoods are equal up to a constant of proportionality (equivalence class of fns).

Q: Does that not affect the calculation of the posterior distribution?

LW: We write the posterior $p(\theta | X_1, \ldots, X_n) \propto L(\theta) p(\theta)$ to convey that.

- We can be more specific; in which case, we have that

$$p(\theta|X_1,\dots,X_n) = \frac{L(\theta)p(\theta)}{\underbrace{\int L(\theta)p(\theta)\,d\theta}}$$

normalisation constant
needed for exact specification of post.

- when using conjugate priors we don't keep track of normalisation constants.
- LW: can leave minimax question for test prep; and also

Q: On Test II prep, Q12b and Q11c, what is happening with priors?

- improper priors $\pi(\mu) = p(\theta) \propto 1$
- not technically a probability distribution, but when multiplied by likelihood yields a proper posterior distribution that is well-defined
- improper prior is a technical artefact of no primary interest.

Q: what is requested when we are asked to find a limiting distribution?

LW: We have the following tools:- $\hat{\theta}_n$ - MLE

CLT: $\sqrt{n}(\hat{\theta}_n - \theta) \longrightarrow N(0, Var(\theta))$

E.g. MLE for Bernoulli :- $\sqrt{n}(\hat{p}-p) \xrightarrow{d} N(0, p(1-p))$

- we mean convergence to a well-defined distribution, usually modulo centring and multiplying
- smooth functions of sequences of random variables that are asymptotically Normal (i.e. sums of IID r.v.s, sample means, standardised sample means, MLE under regularity conditions)

are also asymptotically Normal via delta method.

Delta Method: $\sqrt{n}(g(\hat{p})-g(p)) \xrightarrow{d} N(0, \sigma^2 g'(\hat{p})^2)$

Directly: $P(X_n \leq t) \xrightarrow{d} F(t)$
(using definitions)

MGFS $\psi_n(t) \longrightarrow \psi(t)$

- LW: some distinct examples:

$X_1, \ldots, X_n \sim \text{uniform}(0, \theta)$

$\hat{\theta}_{MLE} = \max(X_1, \ldots, X_n) = X_{(n)}$

$n(\theta - \hat{\theta}_{MLE,n}) \xrightarrow{d} Exp(\theta)$ using direct definition of convergence in distn

Q: what loss functions occur on test?

- only squared error / L_2 loss will be covered on test.
- Hence risk function will be MSE
- Bayes estimator will be posterior mean:-

$\hat{\theta}_{BAYES} = \mathbb{E}[\theta | X_1, \ldots, X_n]$ (under L_2 loss)

Q: can we do practice Test II Bayes estimator questions?

O/S 1 - saved until the end after you have attempted yourself.

Q: can you go over Q4 of Practice Test II (counterexample convergence)

O/S 2 - saved until end.

Q: can you go over Q12 of Practice Test II

O/S 3 - saved until end.

Q: Q3 Practice Test II

- R is known (is there a distinction between this and nuisance parameter?)