Stat210B: Theoretical Statistics	**Lecture Date: March 15, 2007**

Change of Measure and Contiguity

Lecturer: Michael I. Jordan — *Scribe: Aria Haghighi*

In the last lecture, we discussed *contiguity* of measure as the analogue of absolute continuity for asymptotic statistics. In this lecture, we will use contiguity to establish results change-of-measure results for statistical hypothesis testing. We briefly recall the definition of contiguity here,

Definition 1 (Contiguity)**.** Let Q_n and P_n be sequences of measures. We say that Q_n is contiguous w.r.t to P_n, denoted $Q_n \triangleleft P_n$, if for each sequence of measurable sets A_n,[1] we have that

$$P_n(A_n) \to 0 \Rightarrow Q_n(A_n) \to 0$$

We also showed that $Q_n \triangleleft P_n$ if and only if whenever the Radon-Nikodym derivative, $\frac{dQ_n}{dP_n}$, converges weakly under P_n to a random variable V (i.e.$\frac{dQ_n}{dP_n} \overset{P_n}{\rightsquigarrow} V$), then we have $EV = 1$.[2] We also saw that a distribution being in the Quadratic Mean Derivative (QMD) family implied contiguity for shrinking alternatives in statistical testing. Formally, for QMD families P_θ, we have that $P^n_{\theta_0 + \frac{h}{\sqrt{n}}} \triangleleft P^n_{\theta_0}$ (by Theorem 7.2 in van der Vaart (1998) pg. 94). We now state an important result regarding the joint distribution of test statistics and the likelihood ratio:

Lemma 2 (Theorem 6.6 in van der Vaart (1998) pg. 90)**.** *Let P_n and Q_n be sequences of measures such that $Q_n \triangleleft P_n$. Let X_n be a sequence of test statistic random variables. Suppose that we have,*

$$\left(X_n, \frac{dQ_n}{dP_n}\right) \overset{P_n}{\rightsquigarrow} (X, V)$$

for limiting random variables X and V. Then we have that $L(B) = E\mathbf{1}_B(X)V$ defines a measure. Furthermore, $X_n \overset{Q_n}{\rightsquigarrow} L$.

Proof. By contiguity, we have that $EV = 1$, which in turn implies that L must be a probability measure. Using Portmanteau's lemma and a standard induction over measurable functions gives that $X_n \overset{Q_n}{\rightsquigarrow} L$. □

Typically, we have that (X, V) is bi-variate normal. In this case we have a very appealing result about the asymptotic distribution of the test statistic under Q_n.

Lemma 3 (LeCan's Third Lemma, pg. 90 van der Vaart (1998))**.** *Suppose that*

$$(X_n, \log \frac{dQ_n}{dP_n}) \overset{P_n}{\rightsquigarrow} \mathcal{N}\left(\begin{pmatrix} \mu \\ -\frac{1}{2}\sigma^2 \end{pmatrix}, \begin{pmatrix} \Sigma & \tau \\ \tau^T & \sigma^2 \end{pmatrix} \right),$$

where τ and σ are scalars.[3] *Then we have that*

$$X_n \overset{Q_n}{\rightsquigarrow} \mathcal{N}(\mu + \tau, \Sigma)$$

[1] Where measurable means with respect to the underlying Borel set of Q_n, which may change with n.

[2] Note that by Prohorov's theorem that $\frac{dQ_n}{dP_n}$ has a convergent subsequence so the theorem isn't vacuous.

[3] Note that we have that the mean of $\log \frac{dQ_n}{dP_n}$ must be $-\frac{1}{2}\sigma^2$.

This lemma shows that under the alternative distribution Q_n, the limiting distribution of the test statistic X_n is also normal but has a mean shifted by $\tau = \lim_{n\to\infty} \mathrm{Cov}(X_n, \log \frac{dQ_n}{dP_n})$.

Proof. Suppose that (X, W) be the limiting distribution on the RHS of the above. By the continuous mapping theorem we have that,

$$(X_n, \frac{dQ_n}{dP_n}) \overset{P_n}{\rightsquigarrow} (X, e^W)$$

Since we have that $W \sim \mathcal{N}(-\frac{1}{2}\sigma^2, \sigma^2)$, we have that $Q_n \triangleleft P_n$. We have by theorem 6.4 then, that X_n converges weakly to L under Q_n, where $L = E\mathbf{1}_B(X)e^W$. We are going to determine the distribution of L via it's characteristic function,[4]

$$\begin{aligned}
\int e^{it^T x} dL(x) &= E\left[e^{it^T X + W}\right] \\
&= E\left[e^{it^T X + i(-i)W}\right] \\
&= \exp\left\{it^T\mu - \frac{1}{2}\sigma^2 - \frac{1}{2}(t^T, -i)\begin{pmatrix}\Sigma & \tau \\ \tau^T & \sigma^2\end{pmatrix}\begin{pmatrix} t \\ -i\end{pmatrix}\right\} \\
&= e^{it^T(\mu+\tau) - \frac{1}{2}t^T\Sigma t} \\
&\Rightarrow L \sim \mathcal{N}(\mu + \tau, \Sigma)
\end{aligned}$$

□

where the last line is obtained by recognizing the form of the RHS of the previous equation as the characteristic function of the normal distribution.

Example 4 (Asymptotically Linear Statistics)**.** Suppose that P_θ is a family of QMD measures. We are interested in the asymptotic behavior of $\sqrt{n}(\hat{\theta}_n - \theta_0)$. We will consider the following setting,

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = \frac{1}{\sqrt{n}}\sum_i \psi_{\theta_0}(X_i) + o_P(1)$$

where $\mathrm{Var}_{\theta_0}\psi_{\theta_0}(X) = \tau^2 < \infty$ and $E_{\theta_0}\psi_{\theta_0} = 0$. Furthermore, we assume that under H_0 (i.e. when $\theta = \theta_0$), we have by the CLT that,

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathcal{N}(0, \tau^2)$$

Since P_θ is in the QMD family, we have the following expression,

$$\left(\sqrt{n}(\hat{\theta}_n - \theta_0), \frac{dP_{\theta_0 + \frac{h}{\sqrt{n}}}}{dP_{\theta_0}}\right) = \left(\frac{1}{\sqrt{n}}\sum_i \left[\psi_{\theta_0}(X_i), h^T\dot{\ell}_{\theta_0}(X_i)\right] + \left[0, -\frac{1}{2}h^T I_{\theta_0} h\right] + o_P(1)\right)$$

Using the bivariate CLT, we have that the RHS above converges to a normal distribution where the covariance between $\sqrt{n}(\hat{\theta}_n - \theta_0)$ and $\frac{dP_{\theta_0 + \frac{h}{\sqrt{n}}}}{dP_{\theta_0}}$ is given by $\tau = \mathrm{Cov}_{\theta_0}(\psi_{\theta_0}(X), h^T\dot{\ell}_{\theta_0}(X))$.

Our next example builds upon the previous one:

Example 5 (T-Statitic for Location Families)**.** Suppose that $f(X - \theta)$ is a density for a QMD location family. We are interested in testing $\theta_0 = 0$. We define the t-statitic as,

$$t_n = \sqrt{n}\frac{\overline{X}_n}{S_n} = \sqrt{n}\frac{\overline{X}_n}{\sigma} + o_{P_{\theta_0}}(1)$$

[4] Which uniquely determines a distribution.

where the second equality uses a delta method argument. This yields that the t-statistic is an asymptotic linear statistic as in example 4. We are interested in the behavior of t_n under the alternative $\theta_h = \frac{h}{\sqrt{n}}$. Recall that, $\dot{\ell}_{\theta_0} = -\frac{f'(x)}{f(x)}$. Using example 4 and the fact that $\psi_{\theta_0}(X_i) = \frac{X_i}{\sigma}$, we have that

$$\begin{aligned}\tau &= -\frac{h}{\sigma}\text{Cov}(X_i, \frac{f'_{\theta_0}(X_i)}{f_{\theta_0}(X_i)}) \\ &= -\frac{h}{\sigma}\int x\frac{f'}{f}df = -\frac{h}{\sigma}\int xf'dx \\ &= \frac{h}{\sigma}, \text{ using integration by part.}\end{aligned}$$

We therefore have that under shrinking alternatives, $t_n \overset{\frac{h}{\sqrt{n}}}{\rightsquigarrow} \mathcal{N}(\frac{h}{\sigma}, 1)$.

Example 6 (Sign Test for Location Families)**.** We suppose again that $f(X-\theta)$ is a density for QMD family of distributions. We also suppose that $f(\cdot)$ is continuous at the origin and that $P_{\theta=0}(X>0) = \frac{1}{2}$. We define the sign statistic,

$$s_n = \frac{1}{\sqrt{n}}\sum_i (1_{X>0} - \frac{1}{2})$$

We again suppose we are interested in testing whether $\theta_0 = 0$. Under the alternative hypothesis $\theta_h = \frac{h}{\sqrt{n}}$, we have

$$\begin{aligned}\tau &= -h\text{Cov}_{\theta_0}(1_{X>0}, \frac{f'(X)}{f(X)}) \\ &= -h\int_0^\infty f'(X)dx = hf(0)\end{aligned}$$

Under the alternative hypothesis, the asymptotic distribution of s_n is normal with mean $hf(0)$.

References

van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press, Cambridge.