Homework#4, Statistical Inference II, 2013 Spring

Data $(X_1,..., X_n)$ follows independently and identically a continuous distribution with the c.d.f. F . Derive the distribution of the Kolmogoriv-Smirnov statistic $D_n(F) = \sup|F_n(x) - F(x)|$, where $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le x)$.

Answer:

As discussed in class,

$$D_n(F) = \max\{\, D_n^+(F), D_n^+(F)\,\},$$

where

$$D_n^+(F) = \sup\{F_n(x) - F(x)\} = \max_{0 \le i \le n}\left(\frac{i}{n} - U_{(i)}\right),$$

$$D_n^-(F) = \sup\{F(x) - F_n(x)\} = \max_{0 \le i \le n}\left(U_{(i+1)} - \frac{i}{n}\right),$$

where $0 \equiv U_{(0)} < U_{(1)} < \cdots U_{(n)} < U_{(n+1)} \equiv 1$ are order statistics for the uniform random variables. Then,

$$\begin{aligned}
&\Pr\{\, D_n(F) \le t\,\} = \Pr\{\, D_n^+(F) \le t\,, D_n^1(F) \le t\,\} \\
&= \Pr\left(\frac{i}{n} - U_{(i)} \le t, \quad U_{(i+1)} - \frac{i}{n} \le t, \quad \forall i = 0,\ldots,n\right) \qquad (*) \\
&= \Pr\left(\frac{i}{n} - t \le U_{(i)} \le t + \frac{i-1}{n}, \quad \forall i = 1,\ldots,n\right)
\end{aligned}$$

Now we find $\Pr\{\, D_n(F) \le t\,\}$ in three cases:

Case i) $t \ge 1$: $\Pr\{\, D_n(F) \le t\,\} = 1$ since $D_n(F) \le 1$ by definition.

Case ii) $t \le \dfrac{1}{2n}$: Then, $\dfrac{i}{n} - t < t + \dfrac{i-1}{n}$. Hence, the last equation in (*) is zero.

Case iii) $\dfrac{1}{2n} < t < 1$:

The p.d.f. of $(U_{(1)}, \cdots, U_{(n)})$ is $n!\, I(0 < u_1 < \cdots < u_n < 1)$. It follows that

$$\begin{aligned}
&\Pr\left(\frac{i}{n} - t \le U_{(i)} \le t + \frac{i-1}{n}, \quad \forall i = 1,\ldots,n\right) \\
&= n! \prod_{i=1}^{n} \int_{\max\{\,0,\, \frac{i-1}{n} - t\,\}}^{\min\{\, u_{i+1},\, t + \frac{i-1}{n}\,\}} du_1 \ldots du_n = n! \prod_{j=1}^{n} \int_{\max\{\,0,\, \frac{n-j+1}{n} - t\,\}}^{\min\{\, u_{n-j+2},\, t + \frac{n-j}{n}\,\}} du_1 \ldots du_n,
\end{aligned}$$

where $j = n - i + 1$.

Therefore by Cases (i)-(iii),

$$\Pr\{\, D_n(F) \le t\,\} = \left[\prod_{j=1}^{n} \int_{\max\{\,0,\, \frac{n-j+1}{n} - t\,\}}^{\min\{\, u_{n-j+2},\, t + \frac{n-j}{n}\,\}} du_1 \ldots du_n\right] I\left(\frac{1}{2n} < t < 1\right) + I(t \ge 1)\,. \square$$