Advanced Probability I, 2013 Fall, Final exam

Name:

1. Consider the CUSUM process $S_n = \sum_{i=1}^{n}(X_i - mp)$ for the binomial random variables $X_i, \quad i = 1, 2, ...n \overset{iid}{\sim} \text{Bin(m,}p)$. The process is out-of-control at time n if $\max_{1 \le k \le n} |S_k| > c\sqrt{mnp(1-p)}$ occurs. Choose c such that the probability of the out-of-control is less than α at all n.

2. Demonstrate with 2 examples that Fubini's theorem often simplifies the calculations of some integration.

1)

2)

3. Let $X_i, \quad i = 1, 2, ...n \overset{iid}{\sim}$ cdf F defined on $(\Omega, \mathbf{F}, P)$. Let the ecdf be

$$F_n(x, \omega) = \frac{1}{n}\sum_{i=1}^{n} I_{(-\infty, x]}(X_i(\omega)), \quad x \in R .$$

1) State the Glivenko-Cantelli theorem

2) Define $x_{m,k} = F^{-1}(k/m), \quad m \in \mathrm{N}, \quad 1 \le k \le m$, and

$$D_{m,n}(\omega) = \max_k \{ | F_n(x_{m,k}, \omega) - F(x_{m,k}) | \vee | F_n(x_{m,k}-, \omega) - F(x_{m,k}-) | \} .$$

For fixed *m*, verify that $\lim_n D_{m,n}(\omega) = 0$ with probability one.

3) If $x_{m,k-1} \le x < x_{m,k}$, find the upper and lower bounds for $F_n(x, \omega) - F(x)$ in terms of $D_{m,n}(\omega)$ and *m*. [Hint: $F(x_{m,k}-) - F(x_{m,k-1}) \le 1/m$]

4) Prove the Glivenko-Cantelli theorem.

4. Let $Y_i, \quad i = 1, 2, ... \overset{iid}{\sim} \phi(y)$, where $\phi(y)$ is the pdf of standard normal distribution, and $\Im_n = \sigma(Y_1, ..., Y_n)$. Also, let $Y_i, \quad i = 1, 2, ... \overset{iid}{\sim} \phi(y - \Delta)$ for some $\Delta \neq 0$.

i) Calculate the likelihood ratio (X_n) which is martingale w.r.t. $\Im_n$ under some probability measure.

ii) Prove that $(X_n, \Im_n)$ is martingale by using properties of the normal distribution. [e.g., use the mgf of N(0, 1)].

5. Let $X_1,\ldots,X_n \overset{iid}{\sim} \mathrm{N}(\mu_X, 1)$ and $Y_1,\ldots,Y_n \overset{iid}{\sim} \mathrm{N}(\mu_Y, 1)$, where $X_i \perp Y_i$ for $\forall i = 1,\ldots,n$. Suppose that (μ_X, μ_Y) are unknown. We wish to estimate an unknown parameter $\theta = P(X_1 \le Y_1)$.

1) Find a consistent estimator $\hat{\theta}$ of θ.
2) Prove the consistency of $\hat{\theta}$.
3) Derive the convergence in distribution of $\sqrt{n}(\hat{\theta} - \theta)$.

6. Let X_n and Y_n be sequences of random variables with $X_n \xrightarrow{P} x$ and $Y_n \xrightarrow{P} y$. Let $a \in R$ and $b \in R$ be constants (not necessary positive). Show that $aX_n + bY_n \xrightarrow{P} \theta$ for some $\theta \in R$.

Answer 1:

By the maximal inequality (theorem 22.4),

$$P[\max_{1\le k\le n}|S_k|> c\sqrt{mnp(1-p)}\ ]$$
$$\le \frac{1}{\{\ c\sqrt{mnp(1-p)}\ \}^2}\operatorname{var}(S_n)=\frac{1}{c^2 mnp(1-p)}mnp(1-p)=\frac{1}{c^2}$$

Setting $1/c^2=\alpha$, we have $c=1/\sqrt{\alpha}$.

Answer 3:

1) $\sup_x|F_n(x,\omega)-F(x)|\to 0$ with probability one.

2) By the SLLN, $\lim_n F_n(x_{m,k},\omega)=F(x_{m,k})$ with probability one for fixed $x_{m,k}$.

   By the continuous mapping theorem with $f(x,y)=x\vee y$, $\lim_n\{F_n(x_{m,k},\omega)-F(x_{m,k})|\vee|F_n(x_{m,k}-,\omega)-F(x_{m,k}-)|\}=0$ with probability one.

   Again by the continuous mapping theorem $f(x_1,...,x_m)=\max_k(x_k)$, $\lim_n D_{m,n}(\omega)=0$.

3) Note that $F(x_{m,k}-)-F(x_{m,k-1})\le 1/m$. If $x_{m,k-1}\le x<x_{m,k}$,

$$F_n(x,\omega)\le F_n(x_{m,k}-,\omega)=F_n(x_{m,k}-,\omega)-F(x_{m,k}-)+F(x_{m,k}-)$$
$$\le D_{m,n}(\omega)+F(x_{m,k}-)\le D_{m,n}(\omega)+1/m+F(x_{m,k}-)$$
$$\le D_{m,n}(\omega)+1/m+F(x)$$

Similarly, $F_n(x,\omega)\ge -D_{m,n}(\omega)-1/m+F(x)$. Hence,

$-D_{m,n}(\omega)-1/m\le F_n(x,\omega)-F(x)\le D_{m,n}(\omega)+1/m$.

4) Take limit in the previous inequality (omit).

Answer 4:

$$X_n=\prod_{i=1}^n\frac{\phi(y_i-\Delta)}{\phi(y_i)}=\exp\left\{-\frac{1}{2}\sum_{i=1}^n(y_i-\Delta)^2+\frac{1}{2}\sum_{i=1}^n y_i^{\,2}\right\}$$
$$=\exp\left\{\Delta\sum_{i=1}^n\left(y_i-\frac{\Delta}{2}\right)\right\}$$

$$E[X_{n+1} \mid \Im_n] = E\left[\exp\left\{\Delta \sum_{i=1}^{n+1}\left(y_i - \frac{\Delta}{2}\right)\right\} \middle| \Im_n\right]$$

$$= E\left[\exp\left\{\Delta\left(y_n - \frac{\Delta}{2}\right)\right\}\exp\left\{\Delta\sum_{i=1}^{n}\left(y_i - \frac{\Delta}{2}\right)\right\} \middle| \Im_n\right] = \exp\left\{\Delta\sum_{i=1}^{n}\left(y_i - \frac{\Delta}{2}\right)\right\} E\left[\exp\left\{\Delta\left(y_n - \frac{\Delta}{2}\right)\right\}\right]$$

$$= X_n \exp\left(-\frac{\Delta^2}{2}\right) E[\exp\{\Delta y_n\}] = X_n \exp\left(-\frac{\Delta^2}{2}\right) E[\exp\{\Delta y_n\}]$$

$$= X_n \exp\left(-\frac{\Delta^2}{2}\right)\exp\left(\frac{\Delta^2}{2}\right) \qquad (\text{ mgf of } N(0,1)\ )$$

$$= X_n$$

Answer 5:

1) Since $X_1 - Y_1 \sim N(\mu_X - \mu_Y, 2)$, $\theta = P(X_1 - Y_1 \le 0) = \Phi\left(\frac{\mu_Y - \mu_X}{\sqrt{2}}\right)$.

   Let $\hat{\theta} = \Phi\left(\frac{\bar{Y} - \bar{X}}{\sqrt{2}}\right)$, where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$.

2) Note that $\bar{X} \xrightarrow{P} \mu_X$, $\bar{Y} \xrightarrow{P} \mu_Y$. Slutsky's theorem is that

   a) If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} a$, then $X_n Y_n \xrightarrow{d} aX$

   b) If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} a$, then $X_n + Y_n \xrightarrow{d} a + X$.

   c) If $X_n \xrightarrow{P} a$ and $Y_n \xrightarrow{P} b$, then, $X_n + Y_n \xrightarrow{P} a + b$.

   By a) and c), $\frac{\bar{Y} - \bar{X}}{\sqrt{2}} \xrightarrow{P} \frac{\mu_Y - \mu_X}{\sqrt{2}}$.

   By the continuous mapping theorem with $h(x) = \Phi(x)$,

   $\Phi\left\{\frac{\bar{Y} - \bar{X}}{\sqrt{2}}\right\} \xrightarrow{P} \Phi\left\{\frac{\mu_Y - \mu_X}{\sqrt{2}}\right\}$. Hence, $\hat{\theta} \xrightarrow{P} \theta$.

3) By the CLT,

$$\sqrt{n}\left(\frac{\bar{Y} - \bar{X}}{\sqrt{2}} - \frac{\mu_Y - \mu_X}{\sqrt{2}}\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{Y_i - X_i - (\mu_Y - \mu_X)}{\sqrt{2}} \xrightarrow{d} N(0, \frac{\sigma_X^2 + \sigma_Y^2}{2}).$$

   We apply the delta method with $g'(x) = \phi(x)$. Since

$$\phi\left(\frac{\mu_Y - \mu_X}{\sqrt{2}}\right) = \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(\mu_Y - \mu_X)^2}{4}\right\},$$

$$\sqrt{n}\left(\Phi\left(\frac{\bar{Y} - \bar{X}}{\sqrt{2}}\right) - \Phi\left(\frac{\mu_Y - \mu_X}{\sqrt{2}}\right)\right) \xrightarrow{d} N\left(0, \frac{\sigma_X^2 + \sigma_Y^2}{4\pi^2}\exp\left\{-\frac{(\mu_Y - \mu_X)^2}{2}\right\}\right).$$