14.381 PS #6 SOLUTIONS

FALL 2004
TA: JOSE TESSADA
(tessada@mit.edu)

1. C&B 5.3

Note that $Y_i \sim$ Bernoulli with $p_i = P(X_i \geq \mu) = 1 - F(\mu)$ for each i. Since the Y_i's are iid Bernoulli,

$$\sum_{i=1}^{n} Y_i \sim \text{binomial}(n, p) \quad \text{with } p = 1 - F(\mu)$$

2. C&B 5.8

(a)

$$\frac{1}{2n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i - x_j)^2 = \frac{1}{2n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i - \bar{x} + \bar{x} - x_j)^2$$

$$= \frac{1}{2n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} \left[(x_i - \bar{x})^2 - 2(x_i - \bar{x})(x_j - \bar{x}) + (x_j - \bar{x})^2 \right]$$

$$= \frac{1}{2n(n-1)} \left\{ \sum_{i=1}^{n} n(x_i - \bar{x})^2 - 2\sum_{i=1}^{n}(x_i - \bar{x}) \underbrace{\sum_{j=1}^{n}(x_j - \bar{x})}_{=0} + n\sum_{j=1}^{n}(x_j - \bar{x})^2 \right\}$$

$$= \frac{n}{2n(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2 + \frac{n}{2n(n-1)} \sum_{j=1}^{n} (x_j - \bar{x})^2$$

both sums are exactly the same

$$= \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 = S^2$$

(b) It seems induction is the easiest way to proceed, and to simplify things assume, without loss of generality, $\theta_1 = 0$, so $EX_i = 0$.

(i) Prove the equation for $n = 4$.

$$S^2 = \frac{1}{24} \sum_{i=1}^{4} \sum_{j=1}^{4} (x_i - x_j)^2$$

Need to compute $E[M^2]$ and $E[M]^2$, where M is defined as

$$M = \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i - x_j)^2$$

Now $E(M) = 24\theta_2$, because

$$\sum_{j=1}^{n} (x_i - x_j)^2 = \sum_{j \neq i} (x_i - x_j)^2 \quad \text{which only have 3 terms when } n=4.$$

$$\Rightarrow \quad E(M) = \sum_{i=1}^{n} \sum_{j=1}^{n} E(x_i - x_j)^2 \qquad E[(x_i - x_j)^2] = 2\theta_2 \quad i \neq j$$

$$\Rightarrow \quad n \cdot (n-1) \cdot 2\theta_2 = 24\theta_2 \quad \text{if } n=4.$$

Now we need $E(M^2)$: this expected value contains 256 ($=4^4$) terms; we can reduce the computational requirements by looking at terms that have the same structure:

- 112 terms are zero, whenever $i=j$;
- 96 are of the form $E[(x_i - x_j)^2 (x_i - x_k)^2] = \theta_4 + 3\theta_2^2$
- 24 are of the form $E[(x_i - x_j)^4] = 2(\theta_4 + 3\theta_2^2)$
- 24 are of the form $E[(x_i - x_j)^2 (x_k - x_l)^2] = 4\theta_2^2$

We can now compute

$$Var(S^2) = \frac{1}{24^2} \left[24 \cdot 2(\theta_4 + 3\theta_2^2) + 96(\theta_4 + 3\theta_2^2) + 24 \times 4\theta_4 - (24\theta_2)^2 \right]$$

$$= \frac{1}{4} \left[\theta_4 - \frac{1}{3}\theta_2^2 \right]$$

Assume the formula holds for n, and establish it for $n+1$.

Define S_n^2 as S^2 computed using n observations.

We can rewrite

$$S_{n+1}^2 = \frac{1}{2n(n+1)} \left[\underbrace{\sum_{i=1}^{n} \sum_{j=1}^{n} (X_i - X_j)^2}_{M} + 2 \underbrace{\sum_{k=1}^{n} (X_k - X_{n+1})^2}_{Q} \right]$$

then we can write

$$S_{n+1}^2 = \frac{1}{2n(n+1)} (M + 2Q)$$

$$Var(M) = 4 \cdot n \cdot (n-1)^2 \left[\theta_4 - \frac{n-3}{n-1} \theta_2^2 \right]$$

by induction hypothesis

$$Var(Q) = n(n+1)\theta_4 - n(n-3)\theta_2^2$$

X_k and X_{n+1} are independent

$$Cov(M, Q) = 2n(n-1)\left[\theta_4 - \theta_2^2\right]$$

Hence,

$$Var(S_{n+1}^2) = \frac{1}{4n^2(n+1)^2} \left[Var(M) + 4 Var(Q) + 4 Cov(M, Q) \right]$$

$$= \frac{1}{4n^2(n+1)^2} \left[4n(n-1)^2 \left(\theta_4 - \frac{n-3}{n-1} \theta_2^2 \right) + 4n(n+1)\theta_4 - 4n(n-3)\theta_2^2 + 8n(n-1)(\theta_4 - \theta_2^2) \right]$$

$$= \frac{1}{4n^2(n+1)^2} \left[\theta_4 \left(4n(n-1)^2 + 4n(n+1) + 8n(n-1) \right) - \theta_2^2 \left(4n(n-1)(n-3) + 4n(n-3) + 8n(n-1) \right) \right]$$

$$= \frac{1}{4n^2(n+1)^2} \left(4n^2(n+1)\theta_4 - 4n(n+1)(n-2)\theta_2^2 \right)$$

$$= \frac{1}{n+1} \left(\theta_4 - \frac{(n-2)}{n} \theta_2^2 \right)$$

∎

3.- C&B 5.13

Take the hint given in the book and start working from there in the following way:

$$E(c\sqrt{S^2}) = c\sqrt{\frac{\sigma^2}{n-1}} \cdot E\left(\sqrt{\frac{S^2(n-1)}{\sigma^2}}\right) = c\sqrt{\frac{\sigma^2}{n-1}} \cdot E(\sqrt{q})$$

now rememeber that $\frac{S^2(n-1)}{\sigma^2} \sim \chi^2_{(n-1)}$ $\quad q \sim \chi^2_{(n-1)}$

$$= c\sqrt{\frac{\sigma^2}{n-1}} \int_0^\infty \sqrt{q} \frac{1}{\Gamma(\frac{n-1}{2}) 2^{\frac{n-1}{2}}} q^{(\frac{n-1}{2})-1} e^{-q/2} \, dq$$

$$= c\sqrt{\frac{\sigma^2}{n-1}} \frac{\Gamma(n/2)\sqrt{2}}{\Gamma((n-1)/2)} \underbrace{\int_0^\infty \frac{1}{\Gamma(n/2) 2^{n/2}} q^{n/2-1} e^{-q/2} \, dq}_{=1 \quad \text{pdf of } \chi^2_n}$$

$$= \frac{c\,\sigma\,\Gamma(n/2)\sqrt{2}}{\sqrt{n-1}\,\Gamma((n-1)/2)}$$

$$c\,E(\sqrt{S^2}) = c \cdot \frac{\sigma\,\Gamma(n/2)\sqrt{2}}{\sqrt{n-1}\,\Gamma((n-1)/2)}$$

$$\Rightarrow \quad c = \frac{\sqrt{n-1}\,\Gamma((n-1)/2)}{\Gamma(n/2)\sqrt{2}} \quad //$$

4.- C&B 5.16

X_i, $i = 1, 2, 3$ independent $N(i, i^2)$

(a) χ^2_3: $\quad \sum_{i=1}^{3} \left(\frac{X_i - i}{i}\right)^2 \sim \chi^2_3$

(b) $\left(\frac{X_1 - 1}{1}\right) \Big/ \sqrt{\frac{\sum_{i=2}^{3} \left(\frac{X_i - i}{i}\right)^2}{2}} \sim t_2$ $\quad$ Important: the variable in the numerator cannot be in the denominator!

(c) take the answer in (b) and compute its square

$$\frac{\left(\frac{X_1-1}{1}\right)^2}{\frac{1}{2}\left\{\left(\frac{X_2-2}{2}\right)^2+\left(\frac{X_3-3}{3}\right)^2\right\}} \sim F_{1,2}$$

5.- C&B 5.23
We will use the conditional of $Z|x$ and the marginal of $X=x$.

$$P(Z>z)=\sum_{x=1}^{\infty}P(Z>z|x)\,P(X=x)=\sum_{x=1}^{\infty}P(U_1>z,\dots,U_x>z|x)P(X=x)$$
$$=\sum_{x=1}^{\infty}\left\{\prod_{i=1}P(U_i>z)\right\}\cdot P(X=x)$$

this last step uses the fact that U_i is independent of U_j, $i\neq j$; note that according to the formula X affects the result only through the number of terms that are used.

$$=\sum_{x=1}^{\infty}P(U_i>z)^x\,P(X=x)=\sum_{x=1}^{\infty}(1-z)^x\frac{1}{(e-1)x!}$$
$$=\frac{1}{(e-1)}\sum_{x=1}^{\infty}\frac{(1-z)^x}{x!}=\frac{e^{1-z}-1}{e-1} \qquad 0<z<1$$