Statistics 581, Problem Set 3 Solutions
Wellner; 10/27/2001

1. Ferguson, ACILST, page 18, problem 3: Suppose that X_n is a sequence of random variables with densities f_n and X is a random variable with density f with respect to a common dominating measure μ, and that $f_n(x) \to f(x)$ as $n \to \infty$ for all x. Show that
$$Eg(X_n) \to Eg(X)$$
for all bounded measurable functions g.
Solution: First, note that
$$\|g\|_\infty \equiv \sup_{-\infty<x<\infty} |g(x)| < \infty\,.$$
Hence it follows that
$$\begin{aligned} |Eg(X_n) - Eg(X)| &\leq |\int g(x)f_n(x)d\mu(x) - \int g(x)f(x)d\mu(x) \\ &= |\int g(x)(f_n(x) - f(x))d\mu(x)| \\ &\leq \int |g(x)||f_n(x) - f(x)|\, d\mu(x) \\ &\leq \|g\|_\infty \int |f_n(x) - f(x)|d\mu(x) \\ &\to \|g\|_\infty \cdot 0 = 0 \end{aligned}$$
by Scheffé's lemma.

2. Suppose that Y is a random variable with $E(Y^2) < \infty$.
(a) Show that
$$Var(Y) = E\{Var(Y|X)\} + Var\{E(Y|X)\}\,.$$
(b) Suppose that X and Y are random variables defined on the probability space $(\Omega, \mathcal{A}, P)$, and let $\mathcal{D}$ be the $\sigma-$field generated by X; $\mathcal{D} = \sigma[X] \equiv X^{-1}(\mathcal{B})$ where $\mathcal{B}$ is the Borel sigma-field for the real line R. Let $L_2(\Omega, \mathcal{D}, P)$ be the space of all random variables which are $\mathcal{D}$ measurable and Show that $Y - E(Y|X) \perp L_2(\Omega, \mathcal{D}, P)$ in the sense that
$$E\{(Y - E(Y|X))Z\} = 0$$
for all $Z \in L_2(\Omega, \mathcal{D}, P)$.
[Hint: You may use the following fact: if $Z \in L_2(\Omega, \mathcal{D}, P)$, then $Z = g(X)$ for some measurable function g from R to R with $E\{g^2(X)\} < \infty$.]
(c) Interpret (a) and (b) geometrically.
(d) Suppose that $Y \sim \chi^2_n(\delta)$. Compute $E(Y)$ and $Var(Y)$.
Solution: (a) We compute directly:
$$\begin{aligned} Var(Y) &= E[Y - E(Y)]^2 = E[Y - E(Y|X) + E(Y|X) - E(Y)]^2 \\ &= E[Y - E(Y|X)]^2 + 2E[(Y - E(Y|X)][E(Y|X) - E(Y)] + E[E(Y|X) - E(Y)]^2 \\ &= E\{E\{[Y - E(Y|X)]^2|X\}\} + 0 + Var[E(Y|X)] \\ &= E\{Var[Y|X]\} + Var[E(Y|X)] \end{aligned}$$

since, by computing conditionally,

$$
\begin{aligned}
&E[(Y-E(Y|X)][E(Y|X)-E(X)] = E\{E\{[(Y-E(Y|X)][E(Y|X)-E(X)]|X\}\} \\
&= E\{[E(Y|X)-E(X)]E\{[Y-E(Y|X)]|X\}\} \\
&= E\{[E(Y|X)-E(X)]\{E(Y|X)-E(Y|X)\}\} \\
&= E\{[E(Y|X)-E(X)]\cdot 0\} \\
&= 0\,.
\end{aligned}
$$

(b) By the hint, $Z = g(X)$ for some measurable function g. Then, by computing conditionally much as in (a) we have

$$
\begin{aligned}
E\{(Y-E(Y|X))Z\} &= E\{(Y-E(Y|X))g(X)\} \\
&= E\{E\{(Y-E(Y|X))g(X)|X\}\} \\
&= E\{g(X)E\{(Y-E(Y|X))|X\}\} \\
&= E\{g(X)(E(Y|X)-E(Y|X))\} \\
&= E\{g(X)\cdot 0\} = 0\,.
\end{aligned}
$$

(c) A geometric interpretation of (a) is that $Y - E(Y|X)$ is orthogonal to $E(Y|X) - E(Y)$ in $L_2(\Omega, \mathcal{A}, P) = L_2(P)$, thus the identity in (a) can be intepreted as a "pythagorean theorem". (b) says something a bit stronger: $Y - E(Y|X)$ is orthogonal to the whole space consisting of all square-integrable functions of X. Thus $E(Y|X)$ is the orthogonal projection of Y onto this subspace.
(d) Now $(Y|K) \sim \chi^2_{2K+n}$ where $K \sim$ Poisson$(\delta/2)$, so

$$
E(Y) = E\{E(Y|K)\} = E\{2K+n\} = n + 2(\delta/2) = n + \delta\,.
$$

Furthermore,

$$
\begin{aligned}
Var(Y) &= E\{Var(Y|K)\} + Var\{E(Y|K)\} \\
&= E\{2(2K+n)\} + Var\{2K+n\} \\
&= 4(\delta/2) + 2n + 4(\delta/2) \\
&= 2n + 4\delta\,.
\end{aligned}
$$

3. Suppose that X is a random variable with finite fourth moment; $E|X|^4 < \infty$. Then $\mu_4 = E(X-\mu)^4$ is the fourth central moment of X. The ratio $\mu_4/\sigma^4 \equiv \kappa$ is the *kurtosis* of X (or of the distribution function F of X), and $\gamma_2 \equiv \mu_4/\sigma^4 - 3$ is called the *excess of kurtosis*; note that for any $N(\mu, \sigma^2)$ random variable, $\gamma_2 = 0$. Investigate the value of γ_2 for various classical distributions (t_r, uniform, bernoulli, Poission(λ), ...). How big can γ_2 be? How small can γ_2 be? **Solution:** Note that $\mu_4^{1/4} = \{E(X-\mu)^4\}^{1/4} \geq \{E(X-\mu)^2\}^{1/2} = \sigma$ by Liapunov's inequality. Thus $\mu_4/\sigma^4 \geq 1$ always, or $\gamma_2 \equiv \mu_4/\sigma^4 \geq -2$ with equality if $X = \pm 1$ with probability $1/2$ each: then $\mu = 0$, $\sigma^2 = 1$, $\mu_4 = 1$, and $\gamma_2 = -2$.
For $X \sim N(0,1)$, $\gamma_2 = 0$ since $EX^4 = 3$.
For $X \sim t_r$, $r > 4$, $\gamma_2 = 6/(r-4) \nearrow \infty$ as $r \searrow 4$; $\gamma_2 \searrow 0$ as $r \nearrow \infty$.
For $X \sim$ Gamma(α, β), $\gamma_2 = 6/\alpha \nearrow \infty$ as $\alpha \searrow 0$.

For $X \sim$ Poisson(λ), $\gamma_2 = 1/\lambda \nearrow \infty$ as $\lambda \searrow 0$.
For $X \sim$ Bernoulli(p), $\gamma_2 = (1-p)^2/p + p^2/(1-p) - 3$ which $= -2$ when $p = 1/2$, and $\nearrow \infty$ when $p \to 0, 1$.

4. Suppose that $X, X_1, \ldots, X_n$ are i.i.d. with mean μ, variance σ^2, and $E|X|^4 < \infty$.
(a) Show that the sample variance $S_n^2 = \sum_{i=1}^n (X_i - \overline{X}_n)^2/(n-1)$ satisfies

$$\sqrt{n}(S_n^2 - \sigma^2)/\sqrt{2}\sigma^2 \to_d N(0, 1 + \gamma_2/2).$$

where $\mu_4 \equiv E(X-\mu)^4$ and $\gamma_2 \equiv \mu_4/\sigma^4 - 3$ is called the *excess of kurtosis*.
(b) Suppose that you want to test $H : \sigma \le \sigma_0^2$ versus $K : \sigma^2 > \sigma_0^2$ for σ_0 a fixed number, and you base your test on normal theory, but in fact the X's are *not normal* with $\gamma_2 \ne 0$. What effect does this have on the level (or size or actual type one error) of the normal theory test?
Solution: (a) Since

$$\overline{S}_n^2 \equiv \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X}_n)^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2 - (\overline{X}_n - \mu)^2$$

where $Y_i \equiv (X_i - \mu)^2$ are i.i.d. $(E(Y_1), Var(Y_1)) = (\sigma^2, \mu_4 - \sigma^4)$, it follows that

$$\begin{aligned}
\sqrt{n}(\overline{S}_n^2 - \sigma^2) &= \sqrt{n}(\overline{Y}_n - \sigma^2) - \sqrt{n}(\overline{X}_n - \mu)(\overline{X}_n - \mu) \\
&\to_d N(0, \mu_4 - \sigma^4) - N(0, \sigma^2) \cdot 0 \\
&= N(0, \mu_4 - \sigma^4).
\end{aligned}$$

Hence

$$\sqrt{n}(\overline{S}_n^2 - \sigma^2)/(\sqrt{2}\sigma^2) \to_d N(0, (\mu_4/\sigma^4 - 1)/2) = N(0, (\mu_4/\sigma^4 - 3 + 2)/2) = N(0, 1 + \gamma_2/2).$$

If instead of $\overline{S}_n^2$ we consider the more usual $S_n^2 = (n/(n-1))\overline{S}_n^2$, it is easily seen that

$$\sqrt{n}(S_n^2 - \overline{S}_n^2) = \sqrt{n}\left(\frac{n}{n-1} - 1\right)\overline{S}_n^2 = o(1)O_p(1) = o_p(1).$$

Thus we also have

$$\sqrt{n}(S_n^2 - \sigma^2)/(\sqrt{2}\sigma^2) \to_d N(0, 1 + \gamma_2/2).$$

(b) When the X_i's are normal, $\gamma_2 = 0$ and $(n-1)S_n^2/\sigma_0^2 \sim \chi_{n-1}^2$ when $\sigma = \sigma_0$ is true. Hence the size of the normal theory test when normal theory is true is

$$\begin{aligned}
\alpha &= P_{\sigma_0}((n-1)S_n^2/\sigma_0^2 \ge \chi_{n-1,\alpha}^2) \\
&= P_{\sigma_0}\left(\sqrt{n}(S_n^2/\sigma_0^2 - 1)/\sqrt{2} \ge \sqrt{\frac{n}{2}}\left(\frac{\chi_{n-1,\alpha}^2}{n-1} - 1\right)\right).
\end{aligned}$$

Since $\sqrt{n}(S_n^2/\sigma_0^2 - 1)/\sqrt{2} \to_d N(0,1)$ under normality, this forces

$$\sqrt{\frac{n}{2}}\left(\frac{\chi_{n-1,\alpha}^2}{n-1} - 1\right)) \to z_\alpha.$$

Thus when the X_i's *are not normal* we have

$$\begin{aligned} P_{\sigma_0}((n-1)S_n^2/\sigma_0^2 \geq \chi^2_{n-1,\alpha}) &= P_{\sigma_0}(\sqrt{n}(S_n^2/\sigma_0^2 - 1)/\sqrt{2} \geq \sqrt{\frac{n}{2}}(\frac{\chi^2_{n-1,\alpha}}{n-1} - 1)) \\ &\to P(N(0, 1+\gamma_2/2) \geq z_\alpha) \\ &= P(Z \geq \frac{z_\alpha}{\sqrt{1+\gamma_2/2}}) = 1 - \Phi(\frac{z_\alpha}{\sqrt{1+\gamma_2/2}}). \end{aligned}$$

When $\gamma_2 = -2$, the asymptotic size is 0; when $\gamma_2 = 0$, the asymptotic size is α; when $\gamma_2 = \infty$, the asymptotic size is $1/2$. For $\gamma_2 \in [-2, 0)$ the asymptotic size is $< \alpha$ while for $\gamma_2 \in (0, \infty)$ the asymptotic size is $> \alpha$.

5. Suppose that $X_1, \ldots, X_n$ are independent Poisson(λ) random variables (so $P(X_1 = k) = e^{-\lambda}\lambda^k/k!$, $k = 0, 1, \ldots$).
(a) Show that $\sqrt{n}(\overline{X}_n - \lambda) \to_d N(0, \text{"}something\text{"})$.
(b) Show that the sequence $\{\sqrt{n}|\overline{X}_n - \lambda|\}$ is uniformly integrable and find $\lim_{n\to\infty} E(\sqrt{n}|\overline{X}_n - \lambda|)$.
(c) Let $g(x) = x^\gamma$ for $x \geq 0$ and $0 < \gamma < \infty$. Show that $\sqrt{n}(g(\overline{X}_n) - g(\lambda)) \to_d N(0, V^2)$ and compute V^2 explicitly in terms of λ and γ. For what γ is V^2 constant in λ? Is this the value of γ that makes $g(\overline{X}_n)$ "most nearly normal"? (a) By the univariate CLT, since $E(X_1) = \lambda$ and $\text{Var}(X_1) = \lambda$,

$$\sqrt{n}(\overline{X}_n - \lambda) \to_d Y \sim N(0, \lambda);$$

note that $Y =_d \sqrt{\lambda}Z$ where $Z \sim N(0,1)$.
(b) Let $Y_n = \sqrt{n}(\overline{X}_n - \lambda)$. Then $E(Y_n) = 0$ and $E(Y_n^2) = Var(Y_n) = n(\lambda/n) = \lambda$. Hence $\overline{lim}_n E(Y_n^2) = \lambda < \infty$. But this implies that $\{|Y_n|\}_{n\geq 1}$ is uniformly integrable. Hence $E|Y_n| \to E|Y| = \sqrt{\lambda}E|Z| = \sqrt{2\lambda/\pi}$ since

$$E|Z| = 2\int_0^\infty z\phi(z)dz = 2\int_0^\infty \{-\phi'(z)\}dz = 2\phi(0) = \frac{2}{\sqrt{2\pi}} = \sqrt{\frac{2}{\pi}}.$$

(c) Since $g'(x) = \gamma x^{\gamma-1}$ the g' theorem yields

$$\sqrt{n}(g(\overline{X}_n) - g(\lambda)) \to_d g'(\lambda)Y = \gamma\lambda^{\gamma-1}Y \sim N(0, \gamma^2\lambda^{2\gamma-1}).$$

When $\gamma = 1/2$ this yields

$$\sqrt{n}(\sqrt{\overline{X}_n} - \sqrt{\lambda}) \to_d N(0, 1/4).$$

Thus $g(x) = \sqrt{x}$ is variance stabilizing for Poisson. It is not the transformation which makes $g(\overline{X}_n)$ most nearly normal in terms of skewness; it turns out that with $g(x) = x^{2/3}$ the transformed variable $g(\overline{X}_n)$ has approximate skewness 0; see e.g. Anscombe (1948), *Biometrika* **35**, 246 - 254, or Efron (1982), *Ann. Statist.* **10**, 323 - 339. Of course this transformation is *not* variance stabilizing; $\sqrt{n}(\overline{X}_n^{2/3} - \lambda^{2/3}) \to_d N(0, (4/9)\lambda^{1/3})$ which depends on λ.