Statistics 581, Problem Set 2 Solutions
Wellner; 10/11/2006

1. (a) If $W \sim \chi_2^2 = \text{Gamma}(2/2, 1/2) = \text{Gamma}(1, 1/2)$, find the density function f_W, distribution function F_W, and inverse distribution function F_W^{-1} explicitly.
(b) Suppose that $(X, Y) \sim N_2(0, I)$. Show that R and Θ defined by $R^2 = X^2 + Y^2$ and $\Theta = \arctan(Y/X)$ are independent random variables with $R^2 \sim \chi_2^2$ and $\Theta \sim \text{Uniform}(0, 2\pi)$.
(c) Use the results of (a) and (b) to show (using Theorem 2.3.4) how to use two independent Uniform$(0,1)$ random variables U and V to generate two standard normal random variables.

Solution: (a) If $W \sim \chi_2^2 = \text{Gamma}(1, 1/2)$, the density function is given by $f_W(w) = (1/2)e^{-w/2}1_{[0,\infty)}$; i.e. $W \sim \text{Exponential}(1/2)$. Hence the distribution function is $F_W(w) = 1 - \exp(-w/2)$ for $w \geq 0$, and the inverse distribution function is $F_W^{-1}(u) = -2\log(1-u)$.
(b) The joint density of (X, Y) is given by

$$f_{X,Y}(x,y) = \frac{1}{2\pi}\exp(-(x^2+y^2)/2) \qquad \text{for}(x,y) \in R^2 .$$

Moreover, $X = R\cos(\Theta)$ and $Y = R\sin(\Theta)$. Hence the Jacobian of the transformation is

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \left| \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix} \right| = r\cos^2(\theta) + r\sin^2(\theta) = r .$$

Thus we find that the joint density of (R, Θ) is given by

$$\begin{aligned} f_{R,\Theta}(r,\theta) &= f_{X,Y}(r\cos(\theta), r\sin(\theta)) = \frac{1}{2\pi}\exp(-r^2/2)r \qquad \text{on} \quad (0,\infty)\times[0,2\pi] \\ &= r\exp(-r^2/2)\cdot\frac{1}{2\pi} = f_R(r)f_\Theta(\theta) . \end{aligned}$$

Thus R and Θ are independent with densities $f_R(r) = r\exp(-r^2/2)1_{(0,\infty)}$ and $f_\Theta(\theta) = (2\pi)^{-1}1_{[0,2\pi]}(\theta)$. Note that the distribution function of R is given by

$$F_R(r) = \int_0^r f_R(y)dy = \int_0^r y\exp(-y^2/2)dy = 1 - \exp(-r^2/2) .$$

It follows easily from this that

$$F_{R^2}(x) = P(R^2 \leq x) = P(R \leq \sqrt{x}) = 1 - \exp(-x/2)$$

for $x \in [0, \infty)$; i.e. $R^2 \sim$ Exponential$(1/2) =$ Gamma$(1, 1/2) = \chi^2_2$.
(c) If U and V are independent Uniform$(0,1)$ random variables, we can use the inverse transformation to first obtain

$$R^2 \equiv F^{-1}_{\chi^2_2}(U) = -2\log(1-U) \sim \chi^2_2$$

and

$$\Theta \equiv 2\pi V \sim \text{ Uniform}(0, 2\pi).$$

Note that R^2 and Θ are independent by independence of U and V. Then in view of (b)

$$(X, Y) \equiv (R\cos(\Theta), R\sin(\Theta)) \sim N_2(0, I)\,.$$

2. Suppose that Y is a random variable with $E(Y^2) < \infty$; i.e. $Y \in L_2(P)$.
(a) Show that

$$Var(Y) = E\{Var(Y|X)\} + Var\{E(Y|X)\}\,;$$

i.e.

$$E(Y - EY)^2 = E\{(Y - E(Y|X))^2\} + E\{[E(Y|X) - E(Y)]^2\}\,.$$

(b) Interpret (a) geometrically.
(c) Suppose that $Y \sim \chi^2_n(\delta)$. Compute $E(Y)$ and $Var(Y)$.
Hint: Use $E(Y) = E\{E(Y|X)\}$ and (a).

Solution: (a) We compute directly:

$$\begin{aligned}
Var(Y) &= E[Y - E(Y)]^2 = E[Y - E(Y|X) + E(Y|X) - E(Y)]^2 \\
&= E[Y - E(Y|X)]^2 + 2E[(Y - E(Y|X)][E(Y|X) - E(Y)] \\
&\qquad + E[E(Y|X) - E(Y)]^2 \\
&= E\{E\{[Y - E(Y|X)]^2|X\}\} + 0 + Var[E(Y|X)] \\
&= E\{Var[Y|X]\} + Var[E(Y|X)]
\end{aligned}$$

since, by computing conditionally,

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

(b) 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". Also note that $Y - E(Y|X)$ is orthogonal to any function $g(X)$: much as in the last part of (a)

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

(c) Now $(Y|K) \sim \chi^2_{2K+n}$ where $K \sim \text{Poisson}(\delta/2)$, so

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

Furthermore, using part (a) we get

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

3. Suppose that $X \sim F$ on $R^+ \equiv [0, \infty)$, $Y \sim G$ on R^+, and X and Y are independent random variables. Let $Z = \min\{X, Y\} = X \wedge Y$ and $\Delta = 1\{X \leq Y\}$. (This is *right-censored data*: if we view X as a survival time, and Y as a censoring time, then $Z = X$ when $X \leq Y$, but $Z = Y$ when $X > Y$.)
(a) Find the joint distribution of (Z, Δ) by computing the two sub-distribution functions $F_{uc}(z) \equiv P(Z \leq z, \Delta = 1)$ and $F_c(t) \equiv P(Z \leq z, \Delta = 0)$.
(b) If $X \sim \text{Exponential}(\lambda)$ and $Y \sim \text{Exponential}(\mu)$, show that Z and Δ are independent.
(c) Repeat (a) when Z is replaced by $\tilde{Z} = Y$, but Δ is still $\Delta =$

$1\{X \le Y\}$. Are $\tilde{Z}$ and Δ independent when $X \sim \text{Exponential}(\lambda)$ and $Y \sim \text{Exponential}(\mu)$?

Solution: (a) Since $Z = \min\{X, Y\} = X \wedge Y$ and $\Delta = 1\{X \le Y\}$, it follows that

$$H_{uc}(z) \equiv P(X \le z, X \le Y) = \int_{[0,z]} (1 - G(x-))dF(x)\,,$$

and

$$H_c(z) \equiv P(Y \le z, X > Y) = \int_{[0,z]} (1 - F(y))dG(y)\,.$$

These two sub-distribution functions completely determine the joint distribution function H of (Z, Δ) since

$$P(Z \le z, \Delta \le \delta) = \begin{cases} 0, & \text{if } \delta < 0\,, \\ H_c(z), & \text{if } 0 \le \delta < 1\,, \\ H_c(z) + H_{uc}(z), & \text{if } 1 \le \delta < \infty\,. \end{cases}$$

Note that

$$1 - H_c(z) - H_{uc}(z) = P(Z > z) = (1 - F(z))(1 - G(z))\,,$$

so the marginal d.f. of Z is

$$H(z, 1) = H_c(z) + H_{uc}(z) = 1 - (1 - F(z))(1 - G(z))\,.$$

(b) When $1 - F(x) = \exp(-\lambda x)$ and $1 - G(x) = \exp(-\mu x)$, then

$$1 - H(z, 1) = (1 - F(z))(1 - G(z)) = \exp(-(\lambda + \mu)x)\,,$$

while

$$P(\Delta = 1) = P(X \le Y) = H_{uc}(\infty) = \frac{\lambda}{\lambda + \mu}\,,$$

so $Z \sim \text{Exponential}(\lambda + \mu)$, $\Delta \sim \text{Bernoulli}(\lambda/(\lambda + \mu))$. Furthermore,

$$H_{uc}(z) = \int_0^z e^{-\mu x}\lambda e^{-\lambda x}\, dx = \frac{\lambda}{\lambda + \mu}(1 - \exp(-(\lambda + \mu)z))$$

$$H_c(z) = \int_0^z e^{-\lambda x} \lambda e^{-\mu x}\, dx = \frac{\mu}{\lambda+\mu}\left(1 - \exp(-(\lambda+\mu)z)\right) ,$$

so that Z and Δ are independent in this case.
(c) When Z is replaced by $\tilde{Z} = Y$, then since $\tilde{Z} = Y$ and $\Delta = 1\{X \le Y\}$, it follows that

$$\tilde{H}_{uc}(z) \equiv P(\tilde{Z} \le z, \Delta = 1) = P(Y \le z, X \le Y) = \int_{[0,z]} F(y) dG(y) ,$$

and

$$\tilde{H}_c(z) \equiv P(\tilde{Z} \le z, \Delta = 0) = P(Y \le z, X > Y) = \int_{[0,z]} (1 - F(y)) dG(y) .$$

In this case the marginal distributions are given by

$$P(\tilde{Z} > z) = P(Y > z) = \exp(-\mu z),$$
$$P(\Delta = 1) = P(X \le Y) = \frac{\lambda}{\lambda+\mu},$$

but now we have

$$\begin{aligned}
\tilde{H}_c(z) &= P(\tilde{Z} \le z,\, \Delta = 0) = P(Y \le z, X > Y) \\
&= \int_0^z (1 - F(y)) dG(y) = \int_0^z \exp(-\lambda y) \mu e^{-\mu y} dy \\
&= \frac{\mu}{\lambda+\mu}(1 - \exp(-(\lambda+\mu)z)) \\
&\ne P(\Delta = 0) P(\tilde{Z} \le z)
\end{aligned}$$

so $\tilde{Z}$ and Δ are *not* independent for exponential X and Y.

4. (i) Ferguson, ACILST, #6, page 12.
(ii) Show that the problem #6 in Ferguson continues to hold if f_n and g are densities with respect to an arbitrary fixed dominating measure μ.
(iii) Give examples of densities f_n and g satisfying $f_n(x) \to g(x)$ for all x in the two cases μ =Lebesgue measure on R; μ =counting measure on $\{0, 1, \ldots\}$.

Solution: (i) (a) Let $r_n \equiv g - f_n$. Then

$$\begin{aligned} 0 &= 1 - 1 = \int g(x)dx - \int f_n(x)dx = \int (g(x) - f_n(x))dx \\ &= \int r_n(x)dx = \int r_n^+(x)dx - \int r_n^-(x)dx, \end{aligned}$$

so $\int r_n^+(x)dx = \int r_n^-(x)dx$. Thus

$$\begin{aligned} \int |g(x) - f_n(x)|dx &= \int |r_n(x)|dx = \int (r_n^+(x) + r_n^-(x))dx = 2\int r_n^+(x)dx \\ &\to 0 \end{aligned}$$

by the dominated convergence theorem since $r_n^+(x) = (g(x) - f_n(x))^+ \leq g(x)$ where $\int g(x)dx = 1$ and $r_n^+(x) \to 0$.
(b) Now

$$\begin{aligned} &\sup_A |P(X_n \in A) - P(X \in A)| \\ &= \sup_A |\int_A f_n d\lambda - \int_A g d\lambda \leq \sup_A \int_A |f_n - g|d\lambda \\ &\leq \int |f_n - g|d\lambda \to 0 \end{aligned}$$

by hypothesis. Alternatively, letting $P_n(A) \equiv P(X_n \in A)$, $P = P(X \in A)$,

$$d_{TV}(P_n, P) = \frac{1}{2}\int |f_n(x) - g(x)|dx$$

by Proposition 1.1.13 (and our discussion in class) where the right side converges to 0.
(ii) This follows from Propositions 1.1.13 and 1.1.14.
(iii) The two examples given in class on 11/09/06 fill the bill. Alternatively, for the case of dominating measure equal to Lebesgue measure, suppose that $X_n \sim N(\mu_n, \sigma_n^2)$ where $\mu_n \to \mu \in \mathbb{R}$ and $\sigma_n^2 \to \sigma^2 > 0$. Suppose that $X \sim N(\mu, \sigma^2)$. Then the density of X_n is

$$p_n(x) = \frac{1}{\sigma_n}\phi\left(\frac{x - \mu_n}{\sigma_n}\right) \to \frac{1}{\sigma}\phi\left(\frac{x - \mu}{\sigma}\right) \equiv p(x)$$

for each $x \in \mathbb{R}$. Hence by Scheffé's theorem we conclude that

$$d_{TV}(P_n, P) = \frac{1}{2} \int |p_n(x) - p(x)| dx \to 0.$$

For the case of counting measure on $\{0, 1, \ldots\}$, consider the following modification of the Binomial(n, p_n) distribution with $np_n \to \lambda > 0$ as discussed in class: suppose that $X_{n1}, \ldots, X_{nn}$ are independent Bernoulli (p_{n1}), $\ldots$, Bernoulli (p_{nn}). Then let $S_n = \sum_{i=1}^n X_{ni} \sim P_n$ and let $T_n \equiv \sum_{i=1}^n Y_{ni} \sim Q_n$ where $Y_{n1}, \ldots , Y_{nn}$ are independent Poisson(p_{ni}) random variables, and hence $T_n \sim Q_n$ is Poisson$(\sum_1^n p_{nj})$. Then, as can be shown via the optional problem #8 of problem set #3,

$$d_{TV}(P_n, Q_n) \le \sum_{j=1}^n p_{nj}^2.$$

5. Suppose that $X_1, X_2, \ldots$ are iid Exponential(λ). Let $M_n \equiv \min_{1 \le i \le n} X_i$ and $T_n \equiv \max_{1 \le i \le n} X_i$.
(a) Show that $nM_n \stackrel{d}{=}$ exponential(λ).
(b) Show that $T_n - (1/\lambda) \log n \to_d (1/\lambda)T$ where T has the double exponential extreme value distribution function given by $P(T \le x) = \exp(-\exp(-x))$.
(c) Now suppose that $X_1, \ldots, X_n$ are iid distribution function F satisfying $0 < F'(0) < \infty$; here $F'(0)$ is the right-derivative of F at 0:

$$\lim_{x \searrow 0} \frac{F(x) - F(0)}{x} = F'(0).$$

Show that $nM_n \to_d$ exponential $(F'(0))$.

Solution: (a) Now

$$\begin{aligned} P(nM_n > x) &= P\left(\min_{1 \le k \le n} X_k > \frac{x}{n}\right) \\ &= P\left(X_1 > \frac{x}{n}, \ldots, X_n > \frac{x}{n}\right) \\ &= P\left(X_1 > \frac{x}{n}\right) \cdots P\left(X_n > \frac{x}{n}\right) \\ &= P(X_1 > x/n)^n = \{\exp(-\lambda x/n)\}^n \\ &= \exp(-\lambda x). \end{aligned}$$

Hence $nM_n \stackrel{d}{=}$ exponential(λ).
(b) For the maximum T_n,

$$\begin{aligned} P(T_n - (1/\lambda)logn \le x) &= P(\max_{1\le i\le n} X_i \le x + (1/\lambda)\log n) \\ &= P(X_1 \le x + (1/\lambda)\log n)^n \\ &= (1 - \exp(-\lambda x - \log n))^n \\ &= \left(1 - \frac{\exp(-\lambda x)}{n}\right)^n \to \exp(-\exp(-\lambda x)) \\ &= P(T \le \lambda x) \end{aligned}$$

where T has the extreme value distribution function $P(T \le x) = \exp(-\exp(-x)$ for $x \in \mathbb{R}$.
(c) In this case, the reasoning is much the same in part (a) at the beginning:

$$\begin{aligned} P(nM_n > x) &= P(X_1 > x/n)^n \\ &= \left(1 - \frac{nF(x/n)}{n}\right)^n \end{aligned} \tag{1}$$

Since F has a derivative F' (from the right) at 0 and $F(0) = 0$ we have

$$nF(x/n) = \frac{F(x/n) - F(0)}{x/n} \cdot x \to F'(0)x \qquad \text{as} \;\; n \to \infty.$$

Thus the expression on the right side of (1) converges to $\exp(-F'(0)x)$ as $n \to$ inf; i.e. $nM_n \to$ exponential $(F'(0))$.

6. Let $X_n \sim \chi^2_n$ for $n \ge 1$.
(a) Show that $\sqrt{n/2}(n^{-1}X_n - 1) \to_d N(0,1)$.
(b) Is the asymptotic distribution of $Y_n \equiv (1 - 1/n)X_n$ the same as that of X_n? That is, does $\sqrt{n/2}(n^{-1}Y_n - 1) \to_d N(0,1)$? How about that of $T_n \equiv (1 - 1/\sqrt{n})X_n$?
(c) Show that

$$\frac{\sqrt{n}\left\{\left(\frac{X_n}{n}\right)^{1/3} - \left(1 - \frac{2}{9n}\right)\right\}}{\sqrt{2/9}} \to_d N(0,1)$$

as $n \to \infty$. [In fact, this convergence occurs very rapidly; this is the "Wilson - Hilferty" transformation of a Chi-square random variable.

This is shown very clearly by normal probability plots comparing the approximations in (a) and (c).]

Solution: (a) Since $X_n \sim \chi^2_n$, $X_n \stackrel{d}{=} Z_1^2 + \cdots + Z_n^2$ where $Z_i \sim N(0,1)$ with $E(Z_i^2) = 1$ and $Var(Z_i^2) = E(Z_i^4) - (E(Z_i^2))^2 = 3 - 1 = 2$. Hence by the CLT

$$\sqrt{n/2}\left(\frac{X_n}{n} - 1\right) \to_d N(0,1).$$

Now

$$\begin{aligned}
&\sqrt{n/2}\left(\left(1 - \frac{1}{n}\right)\frac{X_n}{n} - 1\right) \\
&= \left(1 - \frac{1}{n}\right)\sqrt{\frac{n}{2}}\left(\frac{X_n}{n} - 1\right) + \sqrt{\frac{n}{2}}\left(1 - \frac{1}{n} - 1\right) \\
&\to_d N(0,1) - 0 = N(0,1)
\end{aligned}$$

so the asymptotic distribution of $(1 - 1/n)X_n$ is the same as that of X_n. However,

$$\begin{aligned}
&\sqrt{\frac{n}{2}}\left(\left(1 - \frac{1}{\sqrt{n}}\right)\frac{X_n}{n} - 1\right) \\
&= \left(1 - \frac{1}{\sqrt{n}}\right)\sqrt{\frac{n}{2}}\left(\frac{X_n}{n} - 1\right) + \sqrt{\frac{n}{2}}\left(1 - \frac{1}{\sqrt{n}} - 1\right)
\end{aligned}$$

(b) Let $g(x) = 3x^{1/3}$, so that $g'(x) = x^{-2/3}$, $g'(1) = 1$. Then

$$\begin{aligned}
&\frac{\left(\frac{X_n}{n}\right)^{1/3} - \left(1 - \frac{2}{9n}\right)}{\sqrt{\frac{2}{9n}}} \\
&= \sqrt{\frac{n}{2}}\left(3\left(\frac{X_n}{n}\right)^{1/3} - 3\right) + \sqrt{\frac{9n}{2}\frac{2}{9n}} \\
&= \sqrt{\frac{n}{2}}\left(g\left(\frac{X_n}{n}\right) - g(1)\right) + \sqrt{\frac{2}{9n}} \\
&\to_d g'(1)X + 0 \sim N(0,1).
\end{aligned}$$

Plots for Problem 6, parts (a) and (d)

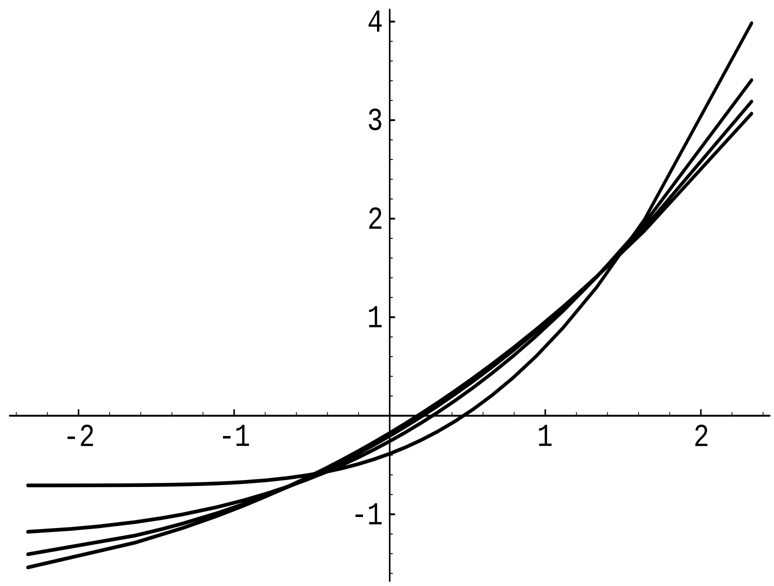

Figure 1: Basic CLT, $n = 3, 5, 7, 9$.

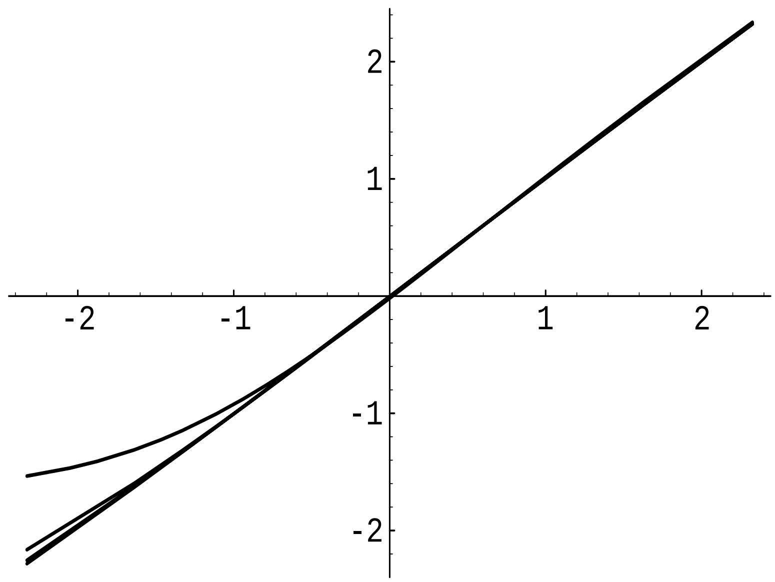

Figure 2: Wilson-Hilferty transform of chi-square, $n = 1, 3, 5, 7$.

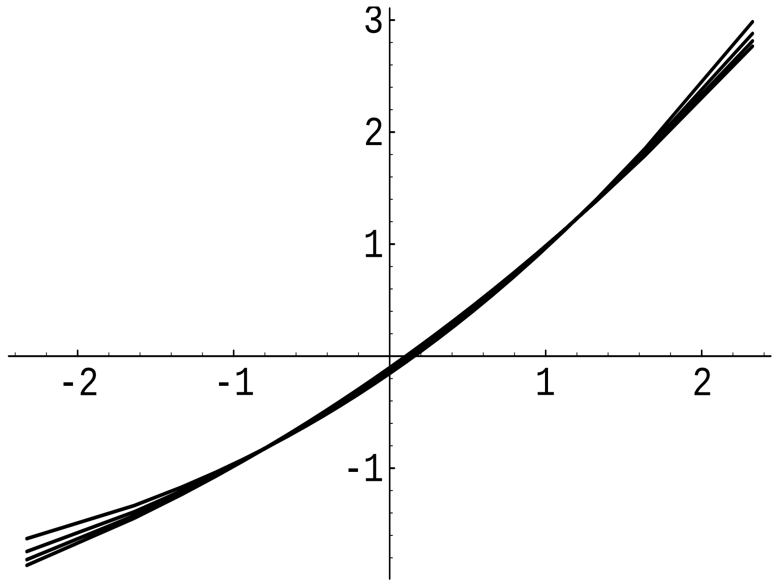

Figure 3: Basic CLT, $n = 9, 13, 17, 21$.

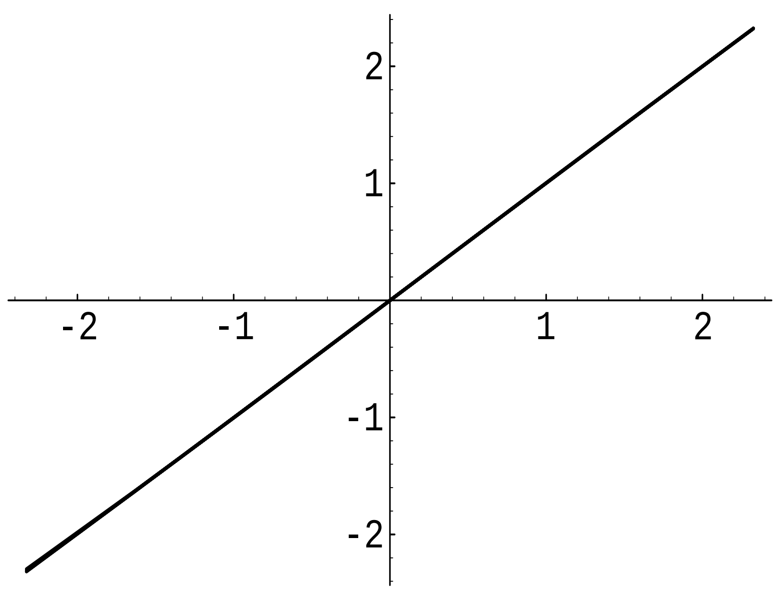

Figure 4: Wilson-Hilferty transform of chi-square,$n = 9, 13, 17, 21$.

Mathematica Code for Figures 1 and 2:

```
<<Statistics`ContinuousDistributions`
dist[j_] := ChiSquareDistribution[j]
gdist := NormalDistribution[0,1]
Qn[u_,j_] := Quantile[dist[j], u]
Tn[u_,j_] := Sqrt[j/2]*(Qn[u,j]/j - 1)
Sn[u_,j_] := Sqrt[9*j/2]*((Qn[u,j]/j)^(1/3) - (1 - 2/(9*j)))
QN[u_] := Quantile[gdist, u]
ParametricPlot[{{QN[u],Tn[u,1]}, {QN[u],Tn[u,3]},
            {QN[u],Tn[u,5]}, {QN[u],Tn[u,7]}}, {u,0.01,.99}]
ParametricPlot[{{QN[u],Sn[u,1]}, {QN[u],Sn[u,3]},
            {QN[u],Sn[u,5]}, {QN[u],Sn[u,7]}}, {u,0.01,.99}]
```