Statistics 581, Problem Set 3, Solutions

Wellner; 10/17/2008. Corrected 10/20/08

1. Ferguson, ACILST, page 34, problem 1(b), modified slightly.
Suppose that $X_1, \ldots, X_n$ is a sample from the Poisson distribution with parameter $\lambda > 0$: $P(X_1 = k) = \exp(-\lambda)\lambda^k/k!$, $k = 0, 1, \ldots$. Let $Z_n = (1/n)\sum_{i=1}^n 1_{[X_i=2]}$.
(a) What is the joint asymptotic distribution of
$$\sqrt{n}((\overline{X}_n, Z_n)' - (\lambda, \lambda^2 e^{-\lambda}/2)')?$$
(b) Let $p_2(\lambda) \equiv P_\lambda(X_1 = 2)$. What is the asymptotic distribution of $\hat{p}_2 \equiv p_2(\hat{\lambda}_n)$ where $\hat{\lambda}_n = \overline{X}_n$?
(c) What is the joint asymptotic distribution of $(Z_n, \hat{p}_2)$ (after centering and rescaling)?
(d) Compute the ratio of the asymptotic variances of the two estimators Z_n and $\hat{p}_2$ of $p_2(\lambda)$. Which estimator would you prefer if the Poisson model (assumption) holds? Which estimator would you prefer if the Poisson model (assumption) fails?

Solution: (a). Let $W_i \equiv (X_i, Y_i) \equiv (X_i, 1_{[X_i=2]})$. Then the W_i's are i.i.d. with mean $E(W_1) = (\lambda, \lambda^2 e^{-\lambda}/2)'$ and covariance matrix
$$\begin{aligned}\Sigma &= \left(\begin{array}{cc} \lambda & 2\lambda^2 e^{-\lambda}/2 - \lambda^3 e^{-\lambda}/2 \\ 2\lambda^2 e^{-\lambda}/2 - \lambda^3 e^{-\lambda}/2 & \frac{\lambda^2 e^{-\lambda}}{2}\left(1 - \frac{\lambda^2 e^{-\lambda}}{2}\right) \end{array}\right) \\ &= \left(\begin{array}{cc} \lambda & \lambda^2(2-\lambda)e^{-\lambda}/2 \\ \lambda^2(2-\lambda)e^{-\lambda}/2 & \frac{\lambda^2 e^{-\lambda}}{2}\left(1 - \frac{\lambda^2 e^{-\lambda}}{2}\right) \end{array}\right). \end{aligned} \tag{1}$$
Hence the multivariate CLT implies that
$$\sqrt{n}(\overline{W} - E(W_1)) = \sqrt{n}((\overline{X}_n, Z_n)' - (\lambda, \lambda^2 e^{-\lambda}/2)) \to_d T \sim N_2(0, \Sigma) \tag{2}$$
where Σ is given in (1).

(b). Now $\hat{p}_2 = g(\overline{X}_n)$ where $g(v) = v^2 e^{-v}/2$. Hence $g'(v) = ve^{-v}(2-v)/2$, $g'(\lambda) = \lambda(2-\lambda)e^{-\lambda}/2$, and $\sqrt{n}(\overline{X}_n - \lambda) \to_d N(0, \lambda)$ by the CLT (or the first component of the convergence in distribution in part (a)). Hence it follows from the delta-method that
$$\sqrt{n}(\hat{p}_2 - p_2(\lambda)) = \sqrt{n}(g(\overline{X}_n) - g(\lambda)) \to_d g'(\lambda)N(0, \lambda) = N(0, \lambda^3(1-\lambda/2)^2 e^{-2\lambda}).$$
(c). At this point it is a bit easier to study $(\hat{p}_2, Z_n) = g(\overline{X}_n, Z_n)$ where $g(u, v) \equiv (u^2 e^{-u}/2, v)$. Then in view of (2) and
$$\nabla g(\lambda, e^{-\lambda}) = \left(\begin{array}{cc} \lambda(1-\lambda/2)e^{-\lambda} & 0 \\ 0 & 1 \end{array}\right),$$

it follows from the delta-method that

$$\sqrt{n}((\hat{p}_2, Z_n)' - (\lambda^2 e^{-\lambda}/2)(1,1)') \to_d \nabla g(\lambda, e^{-\lambda})T \sim N_2(0, \nabla g \Sigma (\nabla g)')$$

where

$$\nabla g \Sigma (\nabla g)' = \begin{pmatrix} \lambda^3(1-\lambda/2)^2 e^{-2\lambda} & \lambda^3(1-\lambda/2)^2 e^{-2\lambda} \\ \lambda^3(1-\lambda/2)^2 e^{-2\lambda} & \frac{\lambda^2 e^{-\lambda}}{2}(1-\frac{\lambda^2 e^{-\lambda}}{2}) \end{pmatrix}.$$

This is a situation in which we have two estimators of $P_\lambda(X_1 = 2) = p_2(\lambda)$, namely the MLE $\hat{p}_2 = p_2(\hat{\lambda})$ and the empirical (or "plug-in" estimator $Z_n = \#\{i \le n : X_i = 2\}/n$. Note that the ratio of the asymptotic variance of $\hat{p}_2$ to the asymptotic variance of Z_n is

$$ARE(\hat{p}_2, Z_n) \equiv \frac{\lambda^3(1-\lambda/2)^2 e^{-2\lambda}}{\frac{\lambda e^{-\lambda}}{2}\left(1-\frac{\lambda e^{-\lambda}}{2}\right)} = \frac{2\lambda(1-\lambda/2)^2 e^{-\lambda}}{(1-\lambda^2 e^{-\lambda}/2)} < 1$$

for all $\lambda > 0$. See the figure below. If the observations really have the assumed Poisson distribution, then the MLE based on the Poisson assumption is preferable because of its smaller asymptotic variance. If the Poisson assumption fails, then the empirical estimator Z_n might be preferable since it always estimates the probability correctly.

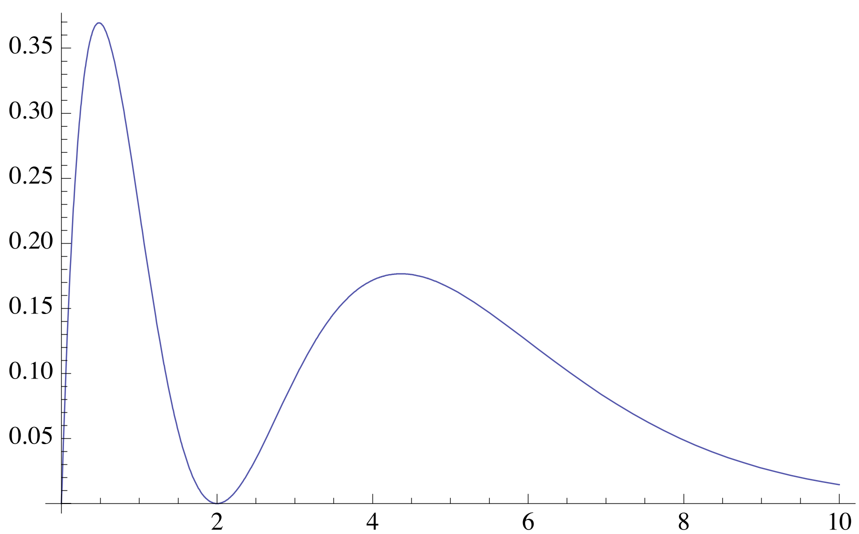

Figure 1: ARE of MLE relative to Plug-In

2. 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 $\mu_4^{1/4} = \{E(X-\mu)^4\}^{1/4} \ge \{E(X-\mu)^2\}^{1/2} = \sigma$ by Liapunov's inequality. Thus $\mu_4/\sigma^4 \ge 1$ always, or $\gamma_2 \equiv \mu_4/\sigma^4 - 3 \ge -2$ 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 $r \nearrow \infty$.
(Note that since $X \stackrel{d}{=} Z/\sqrt{\chi_r^2/r}$ with Z and χ_r^2 independent, it follows that

$$\begin{aligned}
&\mu_4 = E|X|^4 = E|Z|^4 \cdot E(\chi_r^2)^{-2} \cdot r^2 = (3/4)r^2/((r/2-1)(r/2-2)),\\
&\sigma^2 = E|X|^2 = EZ^2 \cdot E(\chi_r^2)^{-1} = (r/2)/(r/2-1),\\
&\gamma_2 = \mu_4/\sigma^4 - 3 = 6/(r-4).)
\end{aligned}$$

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 $p \to 0, 1$.
For $X \sim$ Laplace, or double exponential, with density $2^{-1}\exp(-|x|)$ (as on Ferguson, ACILST page 51), $\mu_4 = 24$ while $\sigma^2 = 2$, so $\gamma_2 = 6$ (not 6.111... as claimed by Ferguson).

3. (From Ferguson, *A Course in Large Sample Theory*, page 65, modified.) In a multinomial experiment with sample size $n = 100$ and 4 cells with null hypothesis $H_0 : \underline{p}_0 = (.2, .3, .3, .2)$, what is the approximate power at the alternative $\underline{p} = (.1, .4, .4, .1)$ when the level of significance is $\alpha = .05$? $\alpha = .01$? How large a sample size is need to achieve power 0.8 at this alternative when $\alpha = .05$? $\alpha = .01$?

   Solution: Now

$$\begin{aligned}
n^{1/2}(\underline{p} - \underline{p}_0) &= 10((.1, .4, .4, .1) - (.2, .3, .3, .2))\\
&= 10(-.1, .1, .1, -.1) = (-1, 1, 1, -1)\,,
\end{aligned}$$

   so the non-centrality parameter is

$$\delta = \frac{1^2}{.2} + \frac{1^2}{.3} + \frac{1^2}{.3} + \frac{1^2}{.2} = \frac{50}{3}\,.$$

Thus the approximate power via $\chi^2_3(\delta)$ is

$$P(\chi^2_3(50/3) \geq \chi_{3,.05}) = P(\chi^2_3(50/3) \geq 7.8147) = .9436\,, \qquad \text{when} \quad \alpha = .05\,,$$

and

$$P(\chi^2_3(50/3) \geq \chi_{3,.01}) = P(\chi^2_3(50/3) \geq 11.345) = .8381 \qquad \text{when} \quad \alpha = .01\,,$$

(b) Now we want to find n so that

$$P(\chi^2_3(\delta_n) \geq 7.81473) = .80$$

where

$$\delta_n = n\left\{\frac{(.1)^2}{.2} + \frac{(.1)^2}{.3} + \frac{(.1)^2}{.3} + \frac{(.1)^2}{.2}\right) = n/6\,.$$

In this case we find that $\delta_n = n/6 = 9.635$, so that $n = 6 \cdot 10.9026 \approx 66$.

When $\alpha = .01$ we find that $\delta_n = n/6 = 15.458$ so that $n = 6 \cdot 15.458 \approx 93$.

The alternative approximation to power that we derived in class is

$$\begin{aligned} P_p(Q_n \geq \chi^2_{k-1,\alpha}) &= P_p\left(\sqrt{n}(n^{-1}Q_n - q) \geq \sqrt{n}(n^{-1}\chi^2_{k-1,\alpha} - q)\right) \\ &\doteq P\left(N(0, d^T A d) \geq \sqrt{n}(n^{-1}\chi^2_{k-1,\alpha} - q)\right) \\ &= 1 - \Phi(\sqrt{n}(n^{-1}\chi^2_{k-1,\alpha} - q)/\sqrt{d^T A d}) \end{aligned}$$

where $d \equiv 2\text{diag}(1/p_0)(p - p_0)$, $A = \text{diag}(p) - pp^T$, and $q = \sum_{j=1}^k (p_j - p_{j0})^2/p_{j0}$. In the present case I calculate $q = 1/6$, $d = (-6, 4, 4, -6)^T/6$, and

$$A = \text{diag}(p) - pp^T = \frac{1}{16}\begin{pmatrix} 9 & -4 & -4 & -1 \\ -4 & 24 & -16 & -4 \\ -4 & -16 & 24 & -4 \\ -1 & -4 & -4 & 9 \end{pmatrix}$$

so that $d^T A d = 6/225$. Thus the approximation becomes

$$P_p(Q_n \geq \chi^2_{3,\alpha}) \doteq 1 - \Phi(\sqrt{n}(n^{-1}\chi^2_{2,\alpha} - 1/6)/\sqrt{61/225}).$$

When I calculate I get

$$\begin{aligned} P_p(Q_n \geq \chi^2_{3,.05}) &\doteq 1 - \Phi(\sqrt{n}(n^{-1}\chi^2_{3,.05} - 1/6)/\sqrt{61/225}) \doteq 0.9554 \\ P_p(Q_n \geq \chi^2_{2,.05}) &\doteq 1 - \Phi(\sqrt{n}(n^{-1}\chi^2_{3,.01} - 1/6)/\sqrt{61/225}) \doteq 0.8466, \end{aligned}$$

which are both somewhat higher than suggested by the non-central chi-square approximation. A Monte-Carlo study not shown here shows that the non-central chi-square approximation is quite accurate in this case. I suspect that the fixed alternative limit theorem and resulting normal approximation to power will do better for more extreme alternatives with a larger number of cells, but I have not carried out a thorough study.

4. Suppose that $X_1, \ldots, X_n$ are independent $N(0,1)$ random variables, and let $Y_i = X_i^2$, for $i = 1, \ldots, n$. Thus $\sum_1^n Y_i \sim \chi_n^2$.
(a) Show that $\sqrt{n}(\overline{Y}_n - 1) \rightarrow_d N(0, \text{“}something''\text{)}$, and find “something”.
(b) Show that for each $r > 0$, $\sqrt{n}(\overline{Y}_n^r - 1) \rightarrow_d N(0, V^2(r))$ and find $V^2(r)$ as a function of r.
(c) Show that

$$\frac{\sqrt{n}(\overline{Y}_n^{1/3} - (1 - 2/(9n)))}{\sqrt{2/9}} \rightarrow_d N(0,1)\,.$$

Does this agree with your result in (b)?
(d) Make normal probablility plots to compare the approximations in (a) and (c). [The transformation in (c) is called the “Wilson-Hilferty” transformation of a χ^2 random variable.

Solution: (a) Since the Y_i's are i.i.d. with $E(Y_i) = 1$ and $Var(Y_i) = E(X_i^4) - E(X_i^2)^2 = 3 - 1 = 2$, it follows from the CLT that

$$\sqrt{n}(\overline{Y}_n - 1) \rightarrow_d Z \sim N(0,2)\,.$$

(b) For $g(x) = x^r$ we have $g'(x) = rx^{r-1}$. Hence by the g'-theorem

$$\begin{aligned} \sqrt{n}(\overline{Y}_n^r - 1) &= \sqrt{n}(g(\overline{Y}_n) - g(1)) \\ &\rightarrow_d g'(1)Z = rN(0,2) = N(0, 2r^2)\,. \end{aligned}$$

Thus $V^2(r) = 2r^2$.
(c) When $r = 1/3$, we find from (b) that

$$\sqrt{n}(\overline{Y}_n^{1/3} - 1) \rightarrow_d (1/3)Z \sim N(0, 2/9)\,.$$

Hence it follows that

$$\begin{aligned} &\sqrt{n}(\overline{Y}_n^{1/3} - (1 - 2/(9n)) \\ &= \sqrt{n}(\overline{Y}_n^{1/3} - 1) + (2/9\sqrt{n}) \\ &\rightarrow_d N(0, 2/9) + 0 = N(0, 2/9)\,. \end{aligned}$$

Hence

$$\frac{\sqrt{n}(\overline{Y}_n^{1/3} - (1 - 2/(9n)))}{\sqrt{2/9}} \rightarrow_d N(0,1)$$

in complete agreement with (b). (The added term $(2/9n)$ gives a higher order approximation to the mean.)
(d) See the plots at the end of this solution set.

5. Suppose that $X_1, X_2, \ldots$ are i.i.d. positive random variables, and define $\overline{X}_n \equiv n^{-1}\sum_{i=1}^n X_i$, $H_n \equiv 1/(n^{-1}\sum_{i=1}^n (1/X_i))$, and $G_n \equiv \{\prod_{i=1}^n X_i\}^{1/n}$ to be the *arithmetic, harmonic, and geometric* means respectively. We know that $\overline{X}_n \rightarrow_{a.s.} E(X_1) = \mu$ if and only if $E|X_1| < \infty$.
(a) Use the SLLN together with appropriate additional hypotheses to show that $H_n \rightarrow_{a.s.} 1/\{E(1/X_1)\} \equiv h$, and $G_n \rightarrow_{a.s.} \exp(E\{\log X_1\}) \equiv g$.
(c) Use the multivariate CLT and the delta method to find the joint limiting distribution of $\sqrt{n}(\overline{X}_n - \mu, H_n - h, G_n - g)$. You will need to impose or assume additional moment conditions to be able to prove this. Specify these additional assumptions carefully.

Solution: (a) If $0 < E(1/X_1) < \infty$, then

$$\frac{1}{n}\sum_{i=1}^n (1/X_i) \rightarrow_{a.s.} E(1/X_1) > 0.$$

If $E|\log(X_1)| < \infty$, then

$$\log G_n = \frac{1}{n}\sum_{i=1}^n \log(X_i) \rightarrow_{a.s.} E\log X_1.$$

Thus by the continuous mapping theorem if both $E(1/X_1) < \infty$ and $E|\log X_1| < \infty$, it follows that

$$(H_n, G_n) \rightarrow_{a.s.} (1/E(1/X_1), \exp(E\log X_1)) \equiv (h, g).$$

(c) By the multivariate CLT, if $EX_1^2 < \infty$, $E(1/X_1)^2 < \infty$, and $E(\log X_1)^2 < \infty$, then

$$\sqrt{n}\begin{pmatrix} \overline{X}_n - \mu \\ \overline{X^{-1}}_n - E(1/X_1) \\ \overline{\log X}_n - E\log X_1 \end{pmatrix} \rightarrow_d \underline{Z} \sim N_3(0, \Sigma)$$

where

$$\Sigma = \begin{pmatrix} Var(X_1) & Cov(X_1, 1/X_1) & Cov(X_1, \log(X_1)) \\ Cov(X_1, 1/X_1) & Var(1/X_1) & Cov(1/X_1, \log X_1) \\ Cov(X_1, \log(X_1)) & Cov(1/X_1, \log X_1) & Var(\log(X_1)) \end{pmatrix}.$$

Hence by the delta method with $g(x, y, z) = (x, 1/y, \exp(z))$ so that $\nabla g(x, y, z) = \mathrm{diag}(1, -y^{-2}, \exp(z))$ and $\nabla g(\mu, E(1/X_1), E(\log X_1)) = \mathrm{diag}(1, -h^2, g)$, it follows that

$$\sqrt{n}\begin{pmatrix} \overline{X}_n - \mu \\ H_n - h \\ G_n - g \end{pmatrix} \to_d \nabla g \cdot \underline{Z} \sim N_3(0, \nabla g \Sigma \nabla g^T).$$

Figures 1 and 2 for problem 4 were generated in Mathematica using the following code:

```
<<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}]
```

Here are the plots for problem 4:

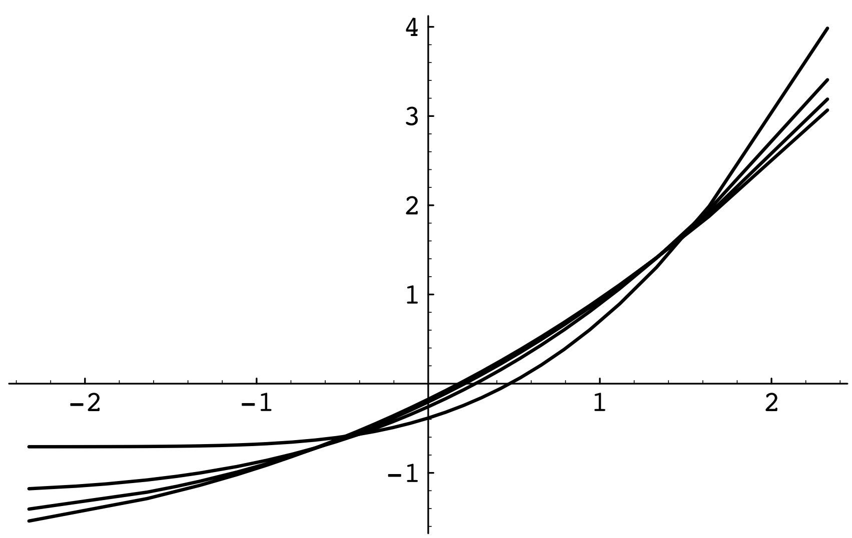

Figure 2: Basic CLT, $n = 3, 5, 7, 9$

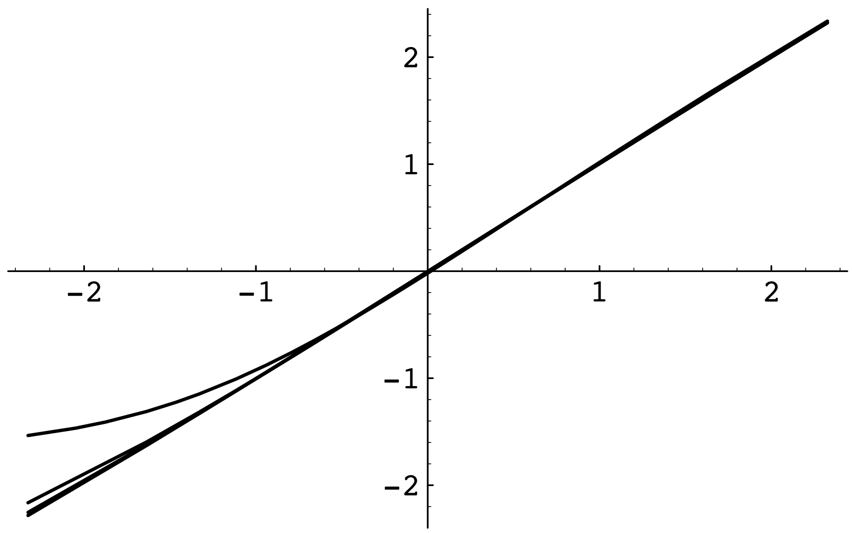

Figure 3: Wilson-Hilferty transform of chi-square, $n = 1, 3, 5, 7$

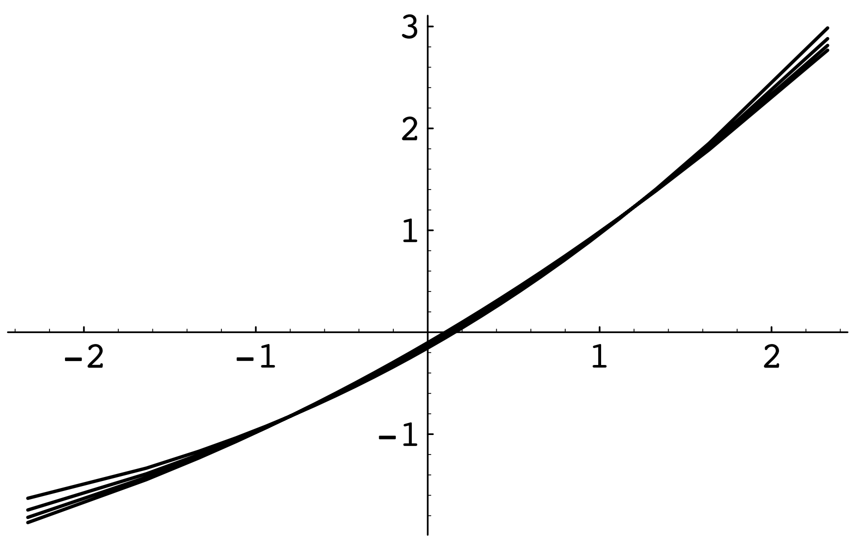

Figure 4: Basic CLT, $n = 9, 13, 17, 21$

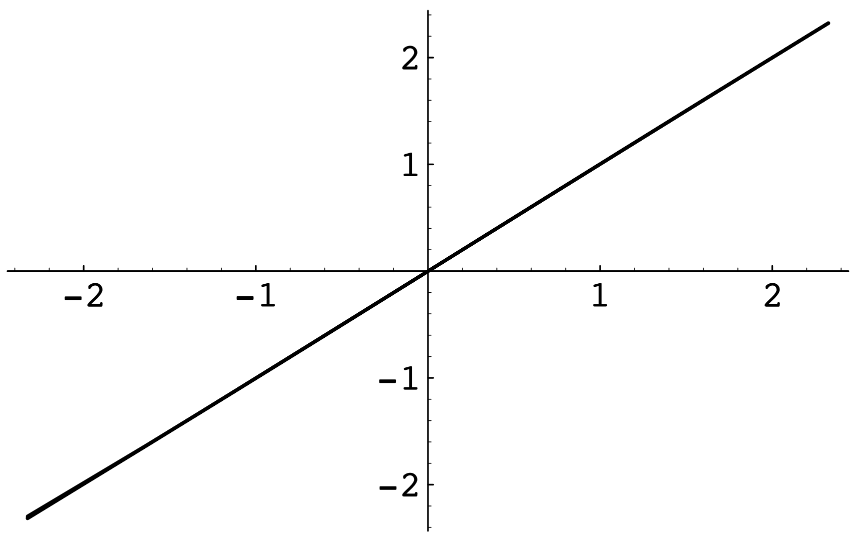

Figure 5: Wilson-Hilferty transform of chi-square, $n = 9, 13, 17, 21$