S. Knoth
Europe University Viadrina

R. Amin
The University of West Florida

Tolerance limits factors for autocorrelated data

- tolerance limits for independent data,
- effects of autocorrelation,
- modified factors.

Tolerance limits

$X \sim \mathcal{N}(\mu, \sigma^2)$

Look for an interval I,
which covers proportion q of the r. v. X:

$$P_X(X \in I) \geq q$$

1. μ, σ^2 known
   $\rightsquigarrow I = \Big(\mu - Z_{(1+q)/2}\,\sigma,\ \mu + Z_{(1+q)/2}\,\sigma\Big)$

2. μ, σ^2 unknown
   $\rightsquigarrow \mu \to \bar{X},\ \sigma \to \hat{\sigma},\ Z_{(1+q)/2} \to ?$

   Now:
   ... interval I, which covers proportion q
   at given confidence level $1 - \alpha$:

$$P_{\bar{X},\hat{\sigma}^2}\Big(P_X\big(\bar{X} - k\,\hat{\sigma} < X < \bar{X} + k\,\hat{\sigma}\big) \geq q\Big) \geq 1 - \alpha$$

Determination of the tolerance limits factor k

iid: $\hat{\sigma}^2 = S^2$.

0. S. S. Wilks (1941/42)

1. A. Wald & J. Wolfowitz (1946):

$$k \approx r\sqrt{\frac{n-1}{\chi^2_{n-1,\alpha}}},$$

$$r : \Phi\left(\frac{1}{\sqrt{n}} + r\right) - \Phi\left(\frac{1}{\sqrt{n}} - r\right) = q\,.$$

2. Improvements/modifications

   Bowker (1946), Ellison (1964), Gardiner & Hull (1966), Howe (1969), Faulkenberry & Daly (1970)

3. Numerical Methods (quadrature)

   Odeh (1978), Odeh/Chou/Owen (1987), Eberhardt/Mee/Reeve (1989)

4. Tables:

   Odeh/Owen (1980) – exact,
   Montgomery (1996) – approximate (Wald/Wolfowitz)

5. Reviews:

   Guenther (1972), Patel (1986)

6. Autocorrelation:

   Amin/Lee (1998/99)

Effects of autocorrelation I

Amin/Lee (1998/99)

estimation of $Var(X)$ – various $\hat{\sigma}^2$

$$\rightarrow S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

$$\text{with } \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \,,$$

$$S_c^2 = \frac{1}{n-1} \sum_{i=2}^{n} (X_i - \hat{\rho}\, X_{i-1})^2$$

$$\text{with } \hat{\rho} = \frac{\sum_{i=2}^{n} (X_i - \bar{X})(X_{i-1} - \bar{X})}{\sum_{i=1}^{n} (X_i - \bar{X})^2} \,,$$

$$\rightarrow \hat{\gamma}_0 = \frac{S_c^2}{1 - \hat{\rho}^2} \,.$$

Effects of autocorrelation II

Amin/Lee (1998/99)

Simulation Study

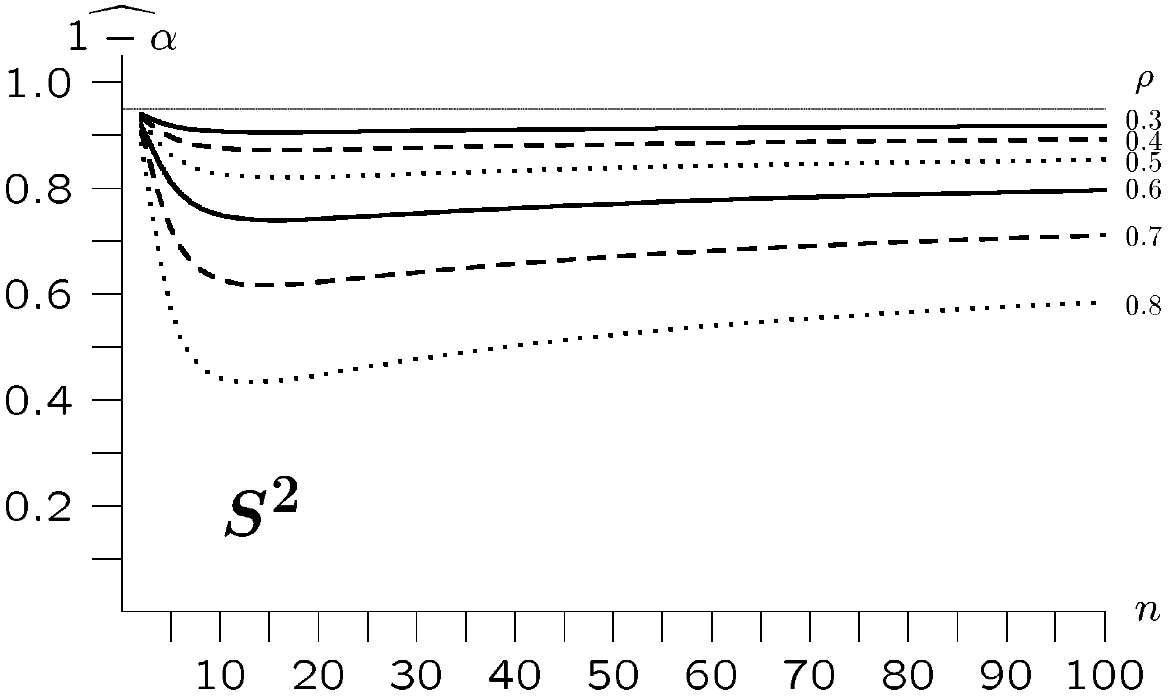
- $k = k(n, q, \alpha)$ as Montgomery (1996),
- simulation of AR(1) data,

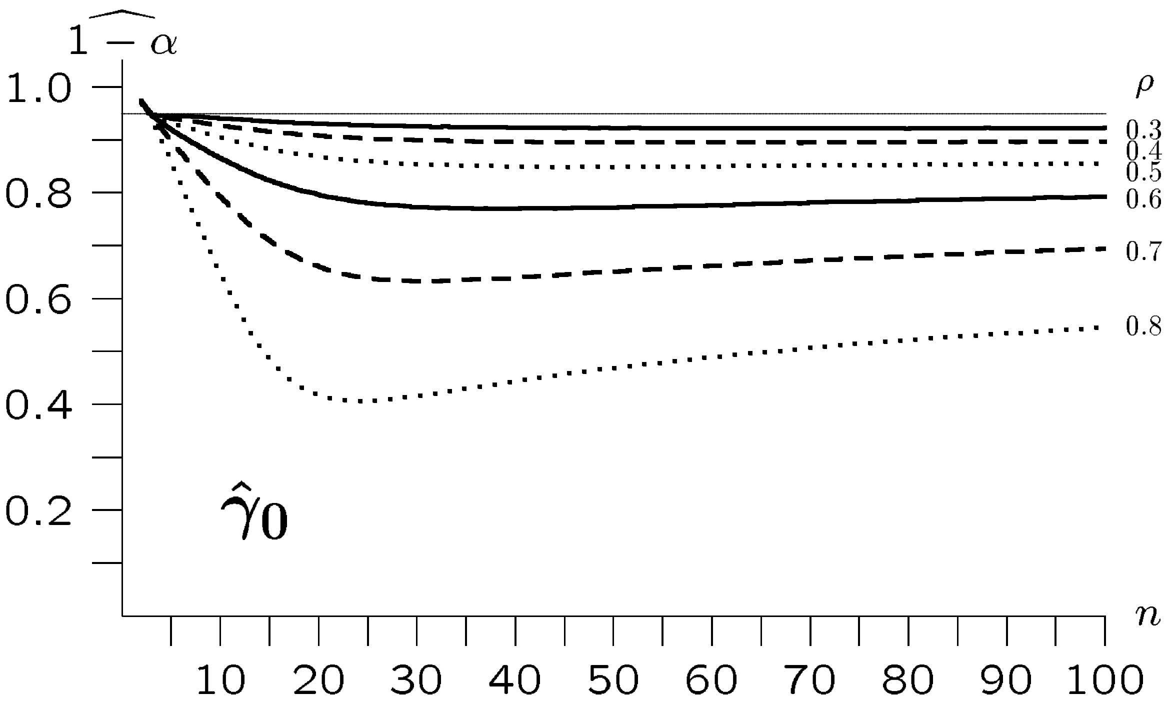
$$\begin{aligned} X_t &= \rho\, X_{t-1} + \varepsilon_t \\ &\quad \varepsilon_t \sim \mathcal{N}(0, \sigma^2 = 1)\ ,\ E(\varepsilon_s\, \varepsilon_t) \underset{s \neq t}{=} 0\,, \end{aligned}$$

- determination of estimators + interval,
- again simulation of AR(1) data,
- relative frequency of
  "interval covers proportion q"
  as empirical significance level,
  comparison with the nominal level $1 - \alpha$.

Actual significance levels

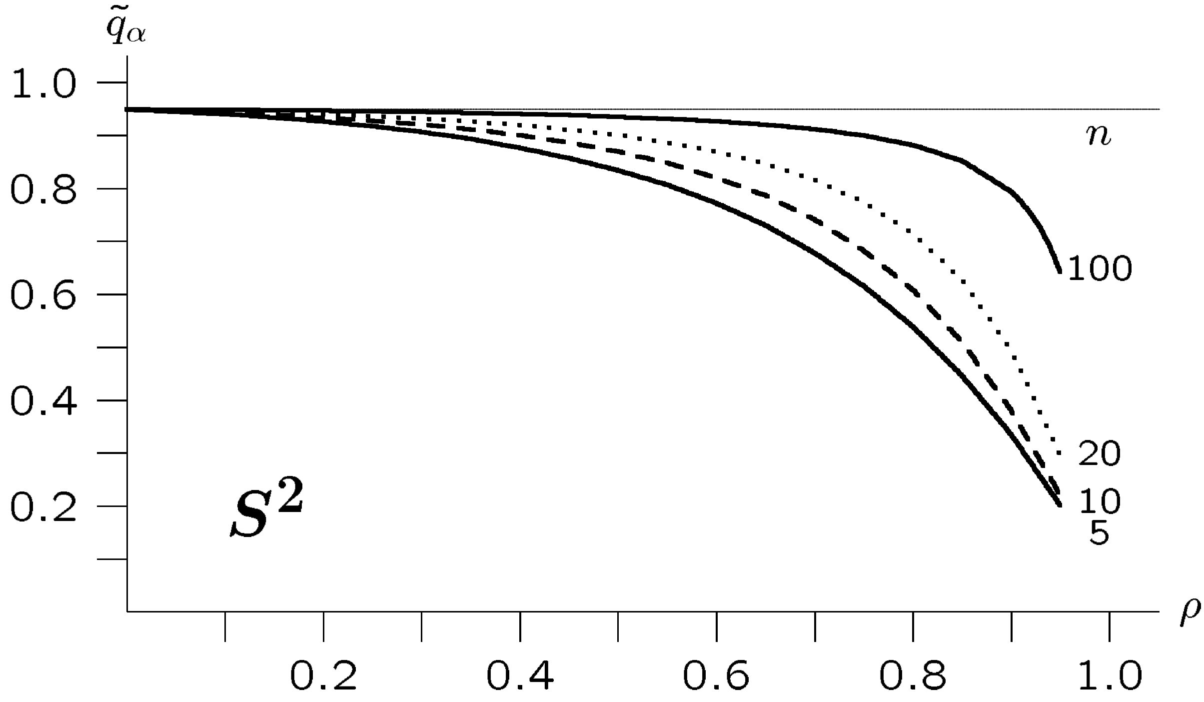
actual coverage $q = 0.95$, nominal $1 - \alpha = 0.95$

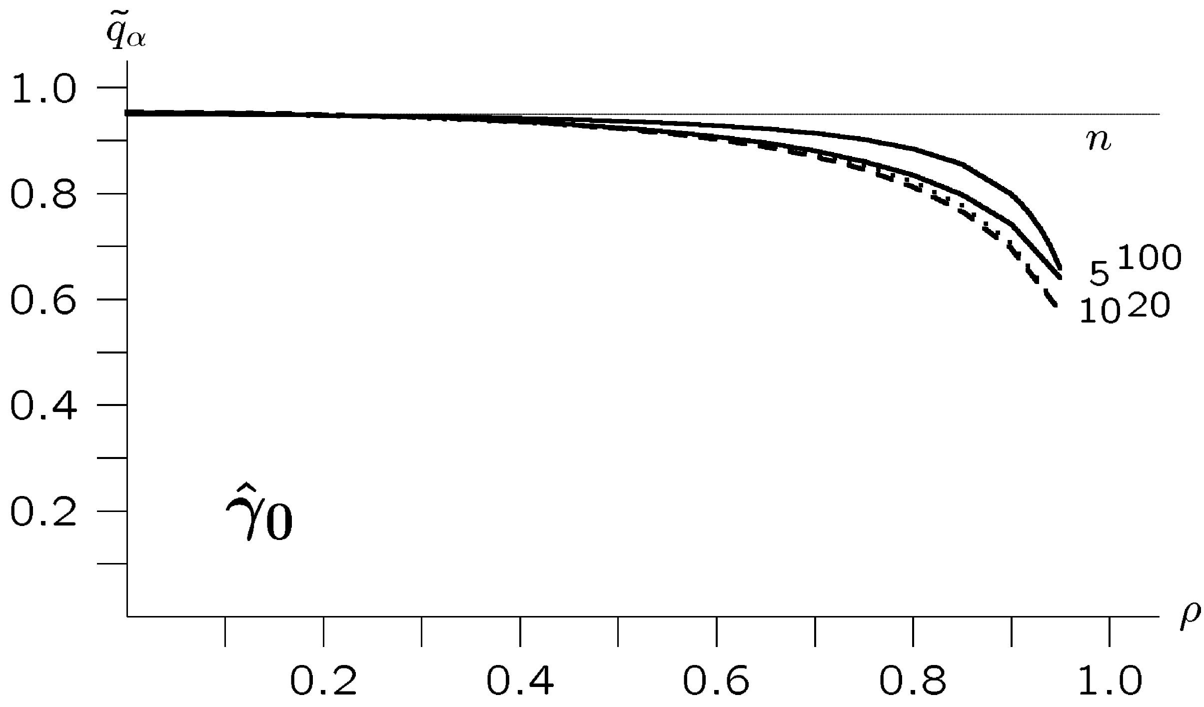




Actual coverage proportion

nominal $q = 0.95$, actual level $1 - \alpha = 0.95$





Modified factors

1. Variance estimators: S^2, $\widehat{\gamma}_0$, S_c^2

2. Computation of the
   actual significance level $\boldsymbol{\eta}(\boldsymbol{k})$
   by using

Schöne/Schmid (1997/99)	or	simulation
S^2 $\leftarrow$	for	$\rightarrow$ $\widehat{\gamma}_0$, S_c^2

Computation of the actual level for S^2

$$\eta(k) = \int_{-\infty}^{\infty}\int_{r^2(x,q)/k^2}^{\infty} f_{\bar{X},S^2}(x,s)\,ds\,dx\,,$$

with $r(x,q)\ :\ \Phi(x+r)-\Phi(x-r) = q$

$$\approx \sum_{i=0}^{m-1}\int_{x_i}^{x_{i+1}}\int_{r_i^2/k^2}^{\infty} f_{\bar{X},S^2}(x,s)\,ds\,dx$$

with $-\infty < x_0 < x_1 < \ldots < x_m < \infty$

and $r_i = r\big((x_i + x_{i+1})/2, q\big)$.

Goodness of approximation

$$\int_{-\infty}^{x_0} \int_{r(x)^2/k^2}^{\infty} f_{\bar{X},S^2}(x,s)\, ds\, dx$$

$$+ \int_{x_m}^{\infty} \int_{r(x)^2/k^2}^{\infty} f_{\bar{X},S^2}(x,s)\, ds\, dx$$

$$< \int_{-\infty}^{x_0} f_{\bar{X}}(x)\, dx + \int_{x_m}^{\infty} f_{\bar{X}}(x)\, dx\, .$$

Here: $x_0 = -x_m$ and

$x_0 \, : \, F_{\bar{X}}(x_0) < 0.5 \cdot 10^{-7}$

$+$ large m (> 1000).

Approximated level for S^2

1. Some arithmetics

$$\begin{aligned}\eta(k) &\approx \sum_{i=0}^{m-1} P\big(\bar{X} \in (x_i, x_{i+1}], S^2 > r_i^2/k^2\big)\\ &= \sum_{i=0}^{m-1} \Big\{F_{\bar{X}}(x_{i+1}) - F_{\bar{X}}(x_i) -\\ &\qquad -\big[F_{\bar{X},S^2}(x_{i+1}, r_i^2/k^2) - F_{\bar{X},S^2}(x_i, r_i^2/k^2)\big]\Big\}.\end{aligned}$$

2. Schöne/Schmid (1997/99)

$$P_{\bar{X},S^2}(|\bar{X}| \le x, S^2 \le s) = \sum_{i,j=0}^{\infty} \frac{f_{i+1,j}}{\sqrt{2\pi}} \frac{4\beta}{2j+1} g_i^{(n+1)}(s)\, x^{2j+1} + + \chi^2_{n-1}\left(\frac{s}{\beta}\right) \left[\Phi\left(\frac{x}{\sqrt{b_0}}\right) - \Phi\left(-\frac{x}{\sqrt{b_0}}\right)\right].$$

Simulation

- $10^6..10^7$ repetitions of $(\bar{X}, \hat{\sigma}^2)$,
- relative frequency of
  $''\,\Phi(\bar{X} + k\,\hat{\sigma}) - \Phi(\bar{X} - k\,\hat{\sigma}) \geq q\,'' \rightsquigarrow \hat{\eta}(k)$.

Determination of k

Give q, α and n.

For a specific variance estimator,
k as numeric solution of

$\eta(k) = 1 - \alpha$ (regula falsi).

Choice of the appropriate variance estimator

Performance measure: average interval width

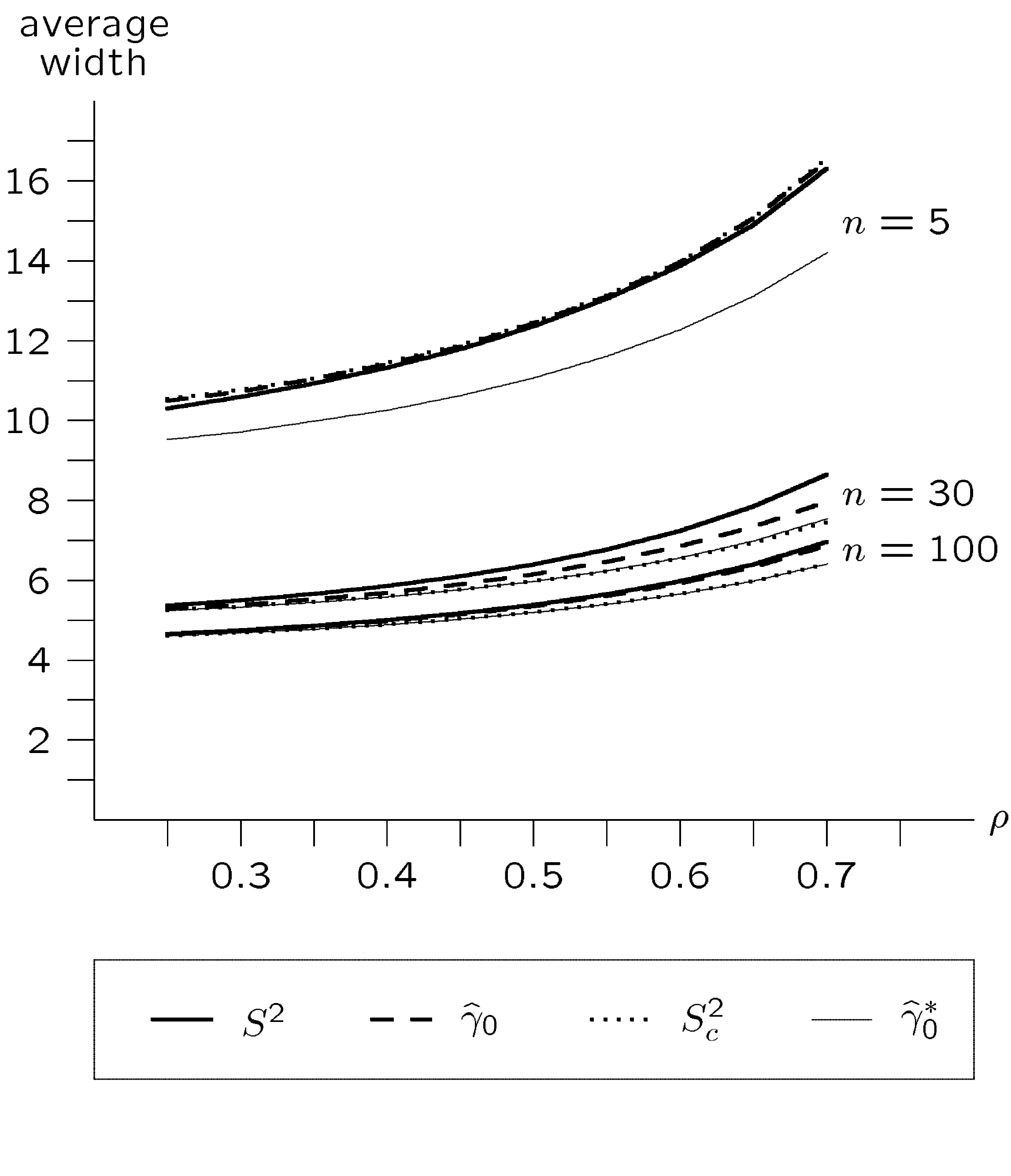
(computation by using simulation)

Benchmark: $\widehat{\gamma}_0^*$

$$\widehat{\gamma}_0^* = \frac{S_c^{*2}}{1-\rho^2},$$
$$S_c^{*2} = \frac{1}{n-1}\sum_{i=2}^{n}(X_i - \rho\, X_{i-1})^2 .$$

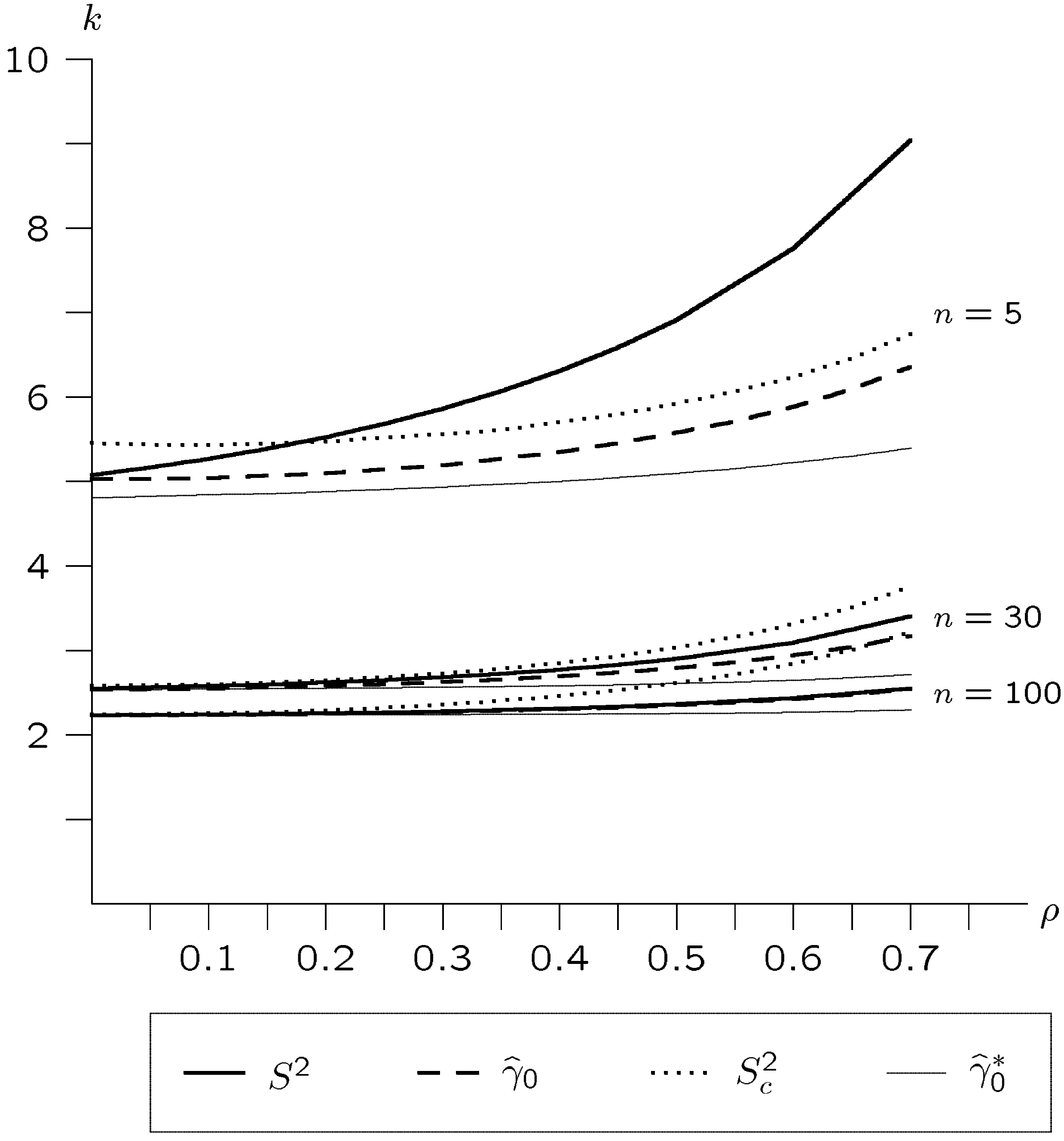
Average interval widths of various tolerance limits

coverage $q = 0.95$, level $1 - \alpha = 0.95$



Robustness of k

coverage $q = 0.95$, level $1 - \alpha = 0.95$



References

R. W. Amin & S. J. Lee (1999) The effects of autocorrelation and outliers on two–sided tolerance limits. *To appear in Journal of Quality Technology.*

D. C. Montgomery (1996) Introduction to Statistical Quality Control. *3rd edition, John Wiley & Sons, New York.*

R. E. Odeh (1978) Tables of two–sided tolerance factors for a normal distribution. *Communications in Statistics – Simulation and Computation B7, 183-201.*

J. K. Patel (1986) Tolerance limits – a review. *Communications in Statistics – Theory and Methods 15, 2719-2762.*

A. Schöne & W. Schmid (1999) On the joint distribution of a quadratic and a linear form in normal variables. *To appear in Journal of Multivariate Analysis.*