Anderson-Darling Tests of Goodness-of-Fit

T. W. Anderson[*]

Stanford University

February 18, 2010

1   Introduction.

A "goodness-of-fit" test is a procedure for determining whether a sample of n observations, $x_1, \ldots, x_n$, can be considered as a sample from a given specified distribution. For example, the distribution might be a normal distribution with mean 0 and variance 1. More generally, the specified distribution is defined as

$$F(x) = \int_{-\infty}^{x} f(y)dy, \quad -\infty < x < \infty, \tag{1}$$

where $f(y)$ is a specified density. This density might be suggested by a theory, or it might be determined by a previous study of similar data.

When X is a random variable with distribution function $F(x) = \Pr\{X \leq x\}$, then $U = F(X)$ is a random variable with distribution function

$$\Pr\{U \leq u\} = \Pr\{F(X) \leq u\} = u, \quad 0 \leq u \leq 1. \tag{2}$$

The model specifies $u_1 = F(x_1), \ldots, u_n = F(x_n)$ as a sample from the distribution (2), that is, the standard uniform distribution on the unit interval $[0, 1]$ written $U(0, 1)$.

A test of the hypothesis that $x_1, \ldots, x_n$ is a sample from a specified distribution, say $F^0(x)$, is equivalent to a test that $u_1 = F^0(x_1), \ldots, u_n = F^0(x_n)$ is a sample from $U(0, 1)$. Define the *empirical distribution function* as

$$F_n(x) = \frac{k}{n}, \quad -\infty < x < \infty, \tag{3}$$

if k of $(x_1, \ldots, x_n)$ are $\leq x$. A goodness-of-fit test is a comparison of $F_n(x)$ with $F^0(x)$. The hypothesis $H_0 : F(x) = F^0(x)$, $-\infty < x < \infty$, is rejected if $F_n(x)$ is very different from $F^0(x)$. "Very different" is defined here as

$$
\begin{aligned}
W_n^2 &= n \int_{-\infty}^{\infty} \left[F_n(x) - F^0(x) \right]^2 \psi \left[F^0(x) \right] dF^0(x) \\
&= n \int_{-\infty}^{\infty} \left[F_n(x) - F^0(x) \right]^2 \psi \left[F^0(x) \right] f^0(x) dx
\end{aligned}
\tag{4}
$$

being large; here (1) holds and $\psi(z)$ is a weight function such that $\psi(z) \geq 0$, and $f^0(x)$ is the density of $F^0(x)$.

If $\psi(z) = 1$, the statistic W_n^2 is the Cramér-von Mises statistic, denoted by $n\omega^2$. Anderson and Darling (1952) gave a table of the limiting distribution of $n\omega^2$ as $n \to \infty$. For example, the 5% significance point is .46136 and the 1% significance point is .74346.

2   The Anderson-Darling Statistic.

For a given x and hypothetical distribution $F^0(\cdot)$, the random variable $n F_n(x)$ has a binomial distribution with probability $F^0(x)$. The expected value of $n F_n(x)$ is $nF^0(x)$ and the variance

is $nF^0(x)\left[1 - F^0(x)\right]$. The definition of the goodness-of-fit statistic (4) permits the choice of weight function $\psi(\cdot)$. In particular the investigator may want to emphasize the tails of the presumed distribution $F^0(x)$. In that case the choice is

$$\psi(u) = \frac{1}{u(1-u)}. \tag{5}$$

Then for a specified x

$$\sqrt{n}\frac{F_n(x) - F^0(x)}{\sqrt{F^0(x)\left[1 - F^0(x)\right]}} \tag{6}$$

has mean 0 and variance 1 when the null hypothesis is true. The Anderson-Darling statistic is

$$A_n^2 = n\int\limits_{-\infty}^{\infty} \frac{\left[F_n(x) - F^0(x)\right]^2}{F^0(x)\left[1 - F^0(x)\right]}\, d\, F^0(x). \tag{7}$$

It was shown in Anderson and Darling (1954) that (7) can be written as

$$A_n^2 = -n - \frac{1}{n}\sum_{j=1}^{n}(2j-1)\left[\log u_{(j)} + \log\left(1 - u_{(n-j+1)}\right)\right] \tag{8}$$

where $u_{(j)} = F^0\left(x_{(j)}\right)$ and $x_{(1)} < x_{(2)} < \ldots < x_{(n)}$ is the ordered sample.

Anderson and Darling found the limiting distribution of A_n^2 [for weight function (5)]. In the next section the development of this distribution is outlined. The 5% significance point of the limiting distribution is 2.492 and the 1% point is 3.880. The mean of this limiting distribution is 1 and the variance is $2(\pi^2 - 9)/3 \sim .57974$.

3   Outline of derivation.

Let $u = F^0(x)$, $u_i = F^{(0)}(x_i)$, $i = 1, \ldots, n$, and $u_{(i)} = F^{(0)}(x_{(i)})$, $i = 1, \ldots, n$. Let $G_n(u)$ be the empirical distribution function of $u_1, \ldots, u_n$; that is,

$$G_n(u) = \frac{k}{n}, \quad 0 \le u \le 1, \tag{9}$$

if k of $u_1, \ldots, u_n$ are $\leq u$. Thus

$$G_n\left[F^0(x)\right] = F_n^0(x), \tag{10}$$

and

$$W_n^2 = n \int_0^1 \left[G_n(u) - u\right]^2 \psi(u)du, \tag{11}$$

when the null hypothesis $F(x) = F^{(0)}(x)$ is true. For every u $(0 \leq u \leq 1)$

$$Y_n(u) = \sqrt{n}\left[G_n(u) - u\right] \tag{12}$$

is a random variable, and the set of these may be considered as a stochastic process with parameter u. Thus

$$\Pr\left\{W_n^2 \leq z\right\} = \Pr\left\{\int_0^1 Y_n^2(u)\psi(u)du \leq z\right\} = A_n(z), \tag{13}$$

say. For a fixed set $u_1, \ldots, u_k$ the k-variate distribution of $Y_n(u_1), \ldots, Y_n(u_k)$ approaches a multivariate normal distribution as $n \to \infty$ with mean and covariance function

$$\mathcal{E}\left[Y_n(u)\right] = 0, \quad \mathcal{E}Y_n(u)Y_n(v) = \min(u, v) - uv. \tag{14}$$

The limiting process of $\{Y_n(u)\}$ is a Gaussian process $y(u)$, $0 \leq u \leq 1$, and $\mathcal{E}y(u) = 0$ and $\mathcal{E}y(u)y(v) = \min(u, v) - uv$. Let

$$a(z) = \Pr\left\{\int_0^1 y^2(u)\psi(u)du \leq z\right\}. \tag{15}$$

Then $A_n(z) \to a(z)$, $0 \leq z < \infty$. The mathematical problem for the Anderson-Darling statistic is to find the distribution function $a(z)$ when $\psi(u) = 1/u(1 - u)$.

We briefly sketch the procedure to find the distribution of $\int_0^1 z^2(u)du$, where $z(u)$ is a Gaussian stochastic process with $\mathcal{E}z(u) = 0$ and $\mathcal{E}z(u)z(v) = k(u, v)$. When the kernel is

4

continuous and square integrable (as is the case here), it can be written as

$$k(u, v) = \sum_{j=1}^{\infty} \frac{1}{\lambda_j} f_j(u) f_j(v), \qquad (16)$$

where λ_j is an eigenvalue and $f_j(u)$ is the corresponding normalized eigenfunction of the integral equation

$$\lambda \int_0^1 k(u, v) f(u) du = f(v), \qquad (17)$$

$$\int_0^1 f_j^2(u) du = 1, \quad \int_0^1 f_i(u) f_j(u) du = 0, \quad i \neq j. \qquad (18)$$

Then the process can be written

$$z(u) = \sum_{j=1}^{\infty} \frac{1}{\sqrt{\lambda_j}} X_j f_j(u), \qquad (19)$$

where $X_1, X_2, \ldots,$ are independent $N(0, 1)$ variables. Then

$$\int_0^1 z^2(u) du = \sum_{j=1}^{\infty} \frac{1}{\lambda_j} X_j^2, \qquad (20)$$

with characteristic function

$$\mathcal{E} \exp \left[it \int_0^1 z^2(u) du \right] = \prod_{j=1}^{\infty} \mathcal{E} \left(\exp it X_j^2 / \lambda_j \right)$$

$$= \prod_{j=1}^{\infty} \left(1 - 2it / \lambda_j \right)^{-\frac{1}{2}}. \qquad (21)$$

The process $Y_n^*(u) = \sqrt{\psi(u)} Y_n(u)$ has covariance function

$$k(u, v) = \sqrt{\psi(u)} \sqrt{\psi(v)} \left[\min(u, v) - uv \right]; \qquad (22)$$

as $n \to \infty$, the process $Y_n^*(u)$ approaches $y^*(u) = \sqrt{\psi(u)} y(u)$ with covariance (22). The characteristic function of the limiting distribution of $n\omega^2$ is

$$\sqrt{\frac{\sqrt{2it}}{\sin \sqrt{2it}}} \qquad (23)$$

5

for $\psi(u) = 1$, and that of the limiting distribution of A_n^2 is

$$\sqrt{\frac{-2\pi it}{\cos\left(\frac{\pi}{2}\sqrt{1+8it}\right)}}. \tag{24}$$

for $\psi(u) = 1/u(1-u)$.

The integral equation (17) can be transformed to a differential equation

$$h''(t) + \lambda\psi(t)\,h((t) = 0. \tag{25}$$

4   Anderson-Darling tests with unknown parameters.

When parameters in the tested distribution are not known, but are estimated efficiently, the covariance (14) is modified, and the subsequent limiting distribution theory for both $n\omega^2$ and A_n^2 follows the same lines as above, with this new covariance. If the parameters are location and/or scale, the limiting distributions do not depend on the true parameter values, but depend on the class of tested distributions. If the parameters are shape parameters, the limiting distribution depends on shape. Limiting distributions have been evaluated and percentage points given for a number of different tested distributions; see Stephens (1976, 1986). Tests for three parameter Weibull, and von Mises have been given by Lockhart and Stephens (1985, 1994).

The percentage points for these tests are much smaller than those given above for the case when parameters are known.

References

Anderson, T.W., and D.A. Darling (1952), "Asymptotic theory of certain 'goodness-of-fit' criteria based on stochastic processes," *Ann. Math, Statistics*, Vol. 23, 193–212.

Anderson, T.W., and D.A. Darling (1954), "A test of goodness-of-fit," *J. Am. Stat. Assoc.*, Vol. 49, 765–769.

Lockhart, R.A. and M.A. Stephens (1985), "Tests of fit for the von-Mises distribution," *Biometrika*, Vol. 72, 647–652.

Lockhart, R.A., and M.A. Stephens (1994), "Estimation and tests of fit for the three–parameter Weibull distribution," *J. Roy. Statist. Soc. B.*, Vol. 56, 491–500.

Stephens, M.A. (1976), "Asymptotic results for goodness-of-fit statistics with unknown parameters," *Ann. Stat.*, Vol. 4, 357–369.

Stephens, M.A. (1986), Chapter 4 in *Goodness-of-fit techniques* (R. D'Agostino and M.A. Stephens, eds.) Marcel Dekker, New York.