Estimation of the Population Prevalence of Adherence to Physical Activity Recommendations based on NHANES Accelerometry Measurements

Kevin W. Dodd, PhD

The observed patterns of NHANES accelerometry data are summarized in Table 1. The Table gives the (population-weighted) fraction of the sample that was observed to have the specified number of “exercise days” out of each possible number of days of wearing the accelerometer.

Table 1. Distribution $P(x,n)$ of observed exercise patterns.

Number of Exercise Days (x)	Number of Days Accelerometer Was Worn (n)						
	1	2	3	4	5	6	7
0	0.049	0.049	0.060	0.076	0.124	0.172	0.248
1	0.002	0.005	0.006	0.007	0.017	0.026	0.038
2	.	0.000	0.001	0.003	0.010	0.014	0.026
3	.	.	0.002	0.001	0.005	0.008	0.012
4	.	.	.	0.002	0.003	0.006	0.011
5	.	.	.	.	0.001	0.003	0.009
6	.	.	.	.	.	0.001	0.003
7	.	.	.	.	.	.	0.001

The physical activity recommendation is to exercise on average 5 or more times per week. We wish to estimate the population prevalence of adherence to this recommendation.

Before any observations are taken, the probability p that a particular individual exercises on a given day is unknown, but is assumed to lie somewhere between 0 and 1. We assume that all possible values of p are equally likely:

$$\pi(p) = 1 \times \mathrm{I}(0 \le p \le 1). \qquad (1)$$

Given p, the number x of exercise days (“successes”) out of n days on which the accelerometer was worn (“trials”) is assumed to be a binomial random variable, say X.:

$$f_X(x \mid n, p) = \Pr(X = x \mid n, p) = \binom{n}{x} p^x (1-p)^{n-x} \text{ for } x = 0, 1, \ldots, n. \qquad (2)$$

Equation (2) describes the probabilistic behavior of the data given the model parameters n and p. However, we wish to describe the probabilistic behavior of the parameter p given the observed data x (and of course the known value of n). By Bayes’ Theorem, the probability density of p given n and X=x is proportional to the product of the “model” shown in Equation (2) and the “prior” shown in Equation (1):

$$h(p \mid n, X = x) \propto \binom{n}{x} p^x (1-p)^{n-x}\, \mathrm{I}(0 \le p \le 1) \tag{3}$$
$$\propto p^x (1-p)^{n-x}\, \mathrm{I}(0 \le p \le 1),$$

where the binomial coefficient is dropped in the last step because it does not depend on *p*. The last expression in Equation (3) is the kernel of a Beta(x+1, n-x+1) density, so the constant of proportionality (required to make the density *h* integrate to unity) is easily obtained from standard statistical texts:

$$h(p \mid n, X = x) = \frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)} p^x (1-p)^{n-x}\, \mathrm{I}(0 \le p \le 1). \tag{4}$$

Therefore, the probability that someone with *x* out of *n* exercise days is adherent (has $p \geq 5/7$) is

$$\Pr(p \ge 5/7 \mid n, X = x) = \int_{5/7}^{1} \frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)} t^x (1-t)^{n-x} dt. \tag{5}$$

Table 2 shows the probability $p_{a|x,n}$=Pr(adhere|x,n) that an individual with *x* out of *n* exercise days is adherent, based on Equation (5).

Table 2. Probability $p_{a|x,n}$ of being adherent given observed exercise pattern.

Number of Exercise Days (x)	Number of Days Accelerometer Was Worn (n)						
	1	2	3	4	5	6	7
0	0.08163	0.02332	0.00666	0.00190	0.00054	0.00016	0.00004
1	0.48980	0.19825	0.07330	0.02570	0.00870	0.00288	0.00093
2	.	0.63557	0.32320	0.14470	0.05970	0.02328	0.00870
3	.	.	0.73969	0.44220	0.22970	0.10827	0.04756
4	.	.	.	0.81407	0.54844	0.32077	0.16899
5	.	.	.	.	0.86719	0.63951	0.41184
6	.	.	.	.	.	0.90514	0.71541
7	.	.	.	.	.	.	0.93224

To obtain the estimated population prevalence of adherence, p_a, we must integrate the conditional probability of adherence given *x* and *n* over the estimated distribution $P(x,n)$ of *n* and *x* in the population. That is,

$$
\begin{aligned}
p_a &= \int_{n,x\in\Omega} p_{a|x,n}\, dP(x,n) \\
&= p_{a|0,1} \Pr(x=0, n=1) \\
&\quad + p_{a|1,1} \Pr(x=1, n=1) \\
&\quad \vdots \\
&\quad + p_{a|7,7} \Pr(x=7, n=7).
\end{aligned} \tag{6}
$$

Note that (6) is simply the weighted average of the adherence probabilities shown in Table 2, where the weights are the estimated population prevalences of each pattern (Table 1). The overall prevalence of adherence, as well as separate estimates of adherence by the number of days the accelerometer was worn, is shown in Table 3.

Table 3. Estimated prevalence of adherence (overall and by number of valid day)

Number of valid days	Estimated prevalence of adherence
Overall	0.032555
1	0.097639
2	0.039521
3	0.038293
4	0.031787
5	0.027959
6	0.026269
7	0.027263