Module 22: Proportions: One Sample

This module presents confidence intervals and tests of hypotheses for proportions for the situation with one random sample from a population.

Proportions

P = Population parameter, the proportion of population with characteristic

x = Number in sample with the characteristic

n = Total number in sample

p = x/n, the sample estimate of the proportion with the characteristic

Hypothesis tests and confidence intervals are based on the normal approximation to the binomial distribution. For the hypothesis:

$H_0: P = P_0$ vs $H_1: P \neq P_0$

The options for the test statistic are:

$$z = \frac{p - P_0}{\sqrt{\frac{p(1-p)}{n}}} \qquad z = \frac{p - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}} \qquad z = \frac{p - P_0}{\sqrt{\frac{P_0 Q_0}{n}}}$$

where $q = 1 - p$

Appointment-Keeping Behavior Re-Evaluated

PHILIP HERTZ, M.Ed., MSPH, AND PAULA L. STAMPS, MS, PhD

Abstract: Many of the traditional approaches to the problem of appointment-keeping behavior have ignored the organizational factors that may be implicated in differentially high broken appointment rates leading to an implicit assumption that low-income and ethnic minority patients will be more likely to break appointments. A case study at a Model Cities Health Center which maintains a kept appointment rate of 85 per cent examined the relationship of broken appointments to age, sex, ethnic background, and payment mechanisms. The results support alternative explanations for differentially high broken appointments centering on the role of the institution in reinforcing appointment-keeping behavior. (Am. J. Public Health 67:1033–1036, 1977)

Broken appointment rates in ambulatory care facilities range from 16 per cent to 40 per cent. It is important to decrease the broken appointment rate as much as possible in order to provide for the best utilization of the direct care staff members. Explanations of broken appointment rates have concentrated primarily upon factors related to the patients, with emphasis placed upon demographic characteristics, such as socioeconomic status, race, age, sex, and educational levels, with other factors such as attitudes toward health care and personality factors also noted. Occasionally, additional factors have also been investigated, including the effect of weather and distance, and variables associated with the organization itself such as the appointment system, staffing patterns, and information flow in the organization. Most studies have concluded with acceptance of a primary myth which is that low-income patients do not keep appointments as well as people in middle class socioeconomic groups. This has been consistently reported by several investigators over the last ten years.[1–7] The conclusions of the relationship between low income and high broken appointments has been proposed for ambulatory care in fee-for-service settings,[1] psychiatric care,[2] and ambulatory care in a pre-paid setting.[3] The magnitude of the differences vary, but in one of the better designed studies, Goodrich contrasted appointment-keeping rates of a pre-paid group and an OEO* Comprehensive Neighborhood Health Care Program that was part of the same medical system. The broken appointment rate for the pre-paid health plan sample was 8 per cent and for the OEO health plan sample 25 per cent.[3]

Confounding variables with low social class include age, ethnic identification, and educational levels. In general, the conclusion of most studies is that the older the patients the more likely they are to keep an appointment,[2, 4] and that members of ethnic minorities, especially Blacks and Spanish-speaking are more likely to break appointments.[2, 4, 5, 7, 8] Patients with lower educational levels are generally concluded to be poorer keepers of appointments,[2] although Adler's study did not confirm this relationship for psychiatric patients.[2] Social disorganization of urban families, especially lower-income families, has also been related to poor appointment keeping behavior.[1]

All of these studies have in common the emphasis upon the failure of the patient to keep an appointment, without analyzing the organizational factors which might be responsible for reinforcing poor appointment-keeping behavior. In those studies where efforts were made to increase the kept appointment rate, some interesting observations have been made with respect to the role of the health facility in broken appointments. Three major factors can be identified as being related to appointment-keeping behavior: physician continuity, communication systems, and type of appointment system. Becker found that physician continuity, i.e., seeing the same physician each time, was positively correlated to appointment-keeping among a group of low-income, mostly Black patients.[1] Increased efforts at communication, including various reminders for missed appointments, also seems to be effective in reducing the broken appointment

*Office of Economic Opportunity.

Address reprint requests to Dr. Paula L. Stamps, Assistant Professor, Health Services Administration, University of Massachusetts, Amherst, MA 01003. Dr. Hertz is Administrator, Holyoke (MA) Health Center. This paper, submitted to the Journal January 18, 1977, was revised and accepted for publication April 21, 1977

patients to see specific requested physicians has been used for some time. Postcard reminders are mailed five days prior to the appointment. In some cases, when the staff deems it necessary, home visits are made to those who break their appointments. Home visits are utilized as many of the patients do not have telephones.

Methodology and Data Collection

Data were collected for the adult medical clinics for a 12-month period of time between April 1, 1974 and March 31, 1975, on number of appointments kept and broken, sex, age, ethnic group, and payment mechanism. Chi-square was utilized at a significance level of 5 per cent to test for the existence of relationships between these variables.

The monthly data for each variable were analyzed. Additionally, a chi-square test on the combined 12 months was calculated. There were a total of 3,172 visits recorded in the adult medical clinic during this time period.

Results

The overall kept appointment rate in the adult medical clinics during the study period was 83 per cent.

Out of the 3,172 adult patient visits, 34 per cent were male and 66 per cent were female. As can be seen from Table 1, there were no significant differences in appointment-keeping behavior between males and females.

The age distribution of the population was broken into two categories using age 44 as the cut-off point. As with other studies,[illegible] the data in Table 1 reveal a tendency for older patients to keep appointments better than younger patients. It should be added that the differences with respect to age are significant for only three months (June, October, and November), but these differences are large enough to contribute to a statistical difference over the entire 12-month period of time.

The majority of the clinic's adult patients are Puerto Rican (69 per cent), with 27 per cent of the patients classified as White and 4 per cent classified as Black. Table 1 also demonstrates that there is a statistical relationship between ethnic group membership and appointment-keeping behavior, with Puerto Ricans having a higher broken appointment rate than either the Whites or Blacks. When these data are analyzed by months, May, June, August, November, February, and March show significant differences. During the months of June and August, neither the appointment secretary nor a bilingual health assistant were present at the Health Center. Both of these persons play key roles in scheduling Spanish-speaking patients. In November, February, and March the adult clinics were changed from their regular night to other evenings and some afternoons in order to accommodate some of the schedules of the physicians. These changes were not adequately conveyed to the non-English speaking patients. The 18.5 per cent broken appointment rate among this group as compared to the 8.5 per cent broken appointment rate among Whites and the 7 per cent broken appointment rate among Blacks appears to be closely associated with these two events. The clinic's failure to fully communicate with the non-English speaking patient population, the erratic staffing of medical clinics, and the changes in the clinic schedules may well have contributed to the higher rates of broken appointments for the Spanish-speaking population. As can be seen from Table 1, these factors did not affect the English-speaking Black and White populations.

Sixty-five per cent of the Health Center's patient population is on Medicaid. As can be seen from Table 1, there is no relationship between Medicaid status and appointment-keeping behavior.

In order to further clarify the relationship of ethnic status and payment mechanism, the three ethnic groups were classified by their payment mechanism. Table 2 shows that there is an association between ethnic group and kept appointments, but this is stronger among those not on Medicaid. This group possesses the same characteristics as the total clinic sample, and thus reflects the same organizational problems that have been previously cited with respect to language problems.

TABLE 1—Distribution of Percentage of Kept and Broken Appointments According to Selected Patient Characteristics

Patient Characteristic	N	% Kept (N = 2,683)	% Not Kept (N = 489)
Sex*			
Male	1,108	85.39	14.61
Female	2,064	84.12	15.88
Age†			
< 45	2,686	83.00	17.00
45 +	486	90.37	9.63
Ethnicity‡			
Puerto Rican	2,206	81.46	18.54
White	847	91.52	8.48
Black	119	95.02	4.98
Payment Mechanism§			
Welfare	2,229	83.40	16.60
Non-Welfare	943	87.38	12.62
TOTAL	3,172	84.58	15.42

*χ^2 = [illegible]
†χ^2 = [illegible]
‡χ^2 = [illegible]
§χ^2 = [illegible]

Summary and Conclusions

The data presented suggest that the difference in the kept appointment rates of different ethnic groups does not reflect a true ethnic difference in health behavior. For the months when there was no significant difference in appointment-keeping behavior between the White and Puerto Rican population, the appointment staff was present, and there were no changes in clinic days or medical staff. During five of the six months when such differences occurred, it was because of the absence of the bilingual appointment staff, the changes made in clinic days, and the physician

22 - 5

TABLE 1—Distribution of Percentage of Kept and Broken Appointments According to Selected Patient Characteristics

Patient Characteristic	N	% Kept (N = 2,683)	% Not Kept (N = 489)
Sex[1]			
Male	1,109	85.39	14.61
Female	2,063	84.15	15.85
Age[2]			
< 45	2,683	83.00	17.00
45+	489	90.37	9.63
Ethnicity[3]			
Puerto Rican	2,206	81.46	18.54
White	837	91.52	8.48
Black	129	93.02	6.98
Payment Mechanism[4]			
Welfare	2,229	83.40	16.60
Non-Welfare	943	87.38	12.62
TOTAL	3,172	84.58	15.42

[1] $X^2 = 8.54$; df = 1; $p > .05$
[2] $X^2 = 24.22$; df = 1; $p < .05$
[3] $X^2 = 53.94$; df = 2; $p < .05$
[4] $X^2 = 8.05$; df = 1; $p > .05$

Example of one sample hypothesis test and confidence interval

Table 1 on the following page indicates that 85.39% of the 1,109 males kept their appointments. Suppose that you had worked in this area previously and the percent of appointments kept before had always been 90% or higher. A question this raises is whether this observed rate is sufficiently different from 90% for you to think your population was different. That is, you would like to test the hypothesis:

$$H_0: P_M = 0.90 \text{ vs } H_1: P_M \neq 0.90$$

Hypothesis test for males keeping appointments

1. *The hypothesis:* H_0: $P_M = 0.90$ vs H_1: $P_M \neq 0.90$

2. *The assumptions:* Independence

3. *The α level:* $\alpha = 0.05$

4. *The test statistic:*

$$z = \frac{p - P_M}{\sqrt{\frac{p(1-p)}{n}}}$$

5. *The rejection region:* Reject H_0: $P_M = 0.90$ if z is not between ± 1.96

6. *The result:* $p = \frac{x}{n} = \frac{947}{1,109} = 0.85$

$$P_M = 0.90$$

$$z = \frac{p - P_M}{\sqrt{\frac{p(1-p)}{n}}}$$

$$= \frac{0.85 - 0.90}{\sqrt{\frac{(0.85)(1-0.85)}{1,109}}}$$

$$= \frac{0.85 - 0.90}{\sqrt{\frac{(0.85)(0.15)}{1,109}}} = \frac{-0.05}{\sqrt{0.000115}} = \frac{-0.05}{0.0107} = -4.66$$

7. *The conclusion:* Reject H_0: $P_M = 0.90$ since z is not between ± 1.96

Confidence Interval for P, with $\alpha = 0.05$

$$C = \left[p - 1.96\sqrt{\frac{p(1-p)}{n}} < P < p + 1.96\sqrt{\frac{p(1-p)}{n}} \right] = 0.95$$

Appointment keeping behavior for males

$$p = \frac{x}{n} = \frac{947}{1{,}109} = 0.85$$

$$C = \left[p - 1.96\sqrt{\frac{p(1-p)}{n}} < P < p + 1.96\sqrt{\frac{p(1-p)}{n}} \right] = 0.95$$

$$C = \left[0.85 - 1.96\sqrt{\frac{(0.85)(0.15)}{1{,}109}} < P < 0.85 + 1.96\sqrt{\frac{(0.85)(0.15)}{1{,}109}} \right] = 0.95$$

Note: $$\sqrt{\frac{(0.85)(0.15)}{1{,}109}} = \sqrt{0.000115} = 0.0107$$

$$C[0.85 - 1.96(0.0107) < P < 0.85 + 1.96(0.0107)] = 0.95$$

$$C[0.85 - 0.021 < P < 0.85 + 0.021] = 0.95$$

$$C[0.829 < P < 0.871] = 0.95$$

Atopy in Children and Parental Social Class

ABSTRACT

[illegible]

Subjects and Methods

Study Population

[illegible]

TABLE 2—Lifetime Prevalence of Atopic Diseases and Characteristic Symptoms as Reported by Parents and Point Prevalence of Physician-Diagnosed Atopic Dermatitis, Allergic Sensitization, and Elevated Total Serum IgE in Children Aged 3 to 14 Years (n = 2408), by Social Class[a] Bad Honnef, Germany, 1992–1993

	Social class[a] I		II		III	
	No.	Prevalence, % (95% CI)	No.	Prevalence, % (95% CI)	No.	Prevalence, % (95% CI)
Atopic disease or symptom (questionnaire information)						
Asthma or wheezy bronchitis[b]	191	4.7 (2.9, 6.5)	1178	5.6 (4.4, 6.8)	1091	10.6 (8.8, 12.5)
Wheezing (more than once)	172	8.7 (4.9, 12.5)	1117	20.9 (18.5, 23.4)	981	23.4 (20.4, 26.5)
Eczema[b]	180	6.2 (3.6, 8.1)	1160	9.6 (8.0, 11.3)	1041	11.8 (9.8, 14.0)
Eczema/skin rashes[b]	169	8.3 (4.8, 11.8)	1139	16.1 (14.0, 18.3)	1021	16.5 (14.2, 18.8)
Allergic rhinitis[b]	182	1.1 (0.1, 3.8)	1169	4.1 (3.0, 5.4)	1034	6.0 (3.7, 8.3)
Allergic reaction[b]	181	10.5 (6.5, 15.8)	1170	12.8 (10.8, 14.7)	1033	15.3 (13.2, 17.7)
Atopic dermatitis (dermatologic examination)	193	6.0 (3.6, 9.2)	1195	6.9 (5.5, 8.6)	999	9.9 (7.9, 12.4)
Allergic sensitization (skin prick test)						
At least 1 of 12 tested allergens	186	11.8 (7.1, 17.4)	1207	20.6 (18.3, 23.2)	994	28.1 (25.3, 31.0)
Mite allergens only (dpt or dpf)	186	6.0 (3.0, 9.9)	1207	9.3 (7.5, 11.5)	994	8.3 (6.6, 10.3)
Pollen allergens only (mugwort, alternaria)	186	5.9 (2.8, 10.7)	1207	8.5 (6.8, 10.5)	994	7.5 (5.9, 9.4)
Animal dander allergens, mold, storage mites, mugwort	186	7.3 (3.8, 12.4)	1207	14.3 (12.2, 16.6)	994	16.6 (13.9, 19.6)
Cat epithelium	186	6.1 (2.8, 9.9)	1207	8.1 (6.6, 9.9)	994	8.7 (6.9, 11.1)
At least 2 of 12 tested allergens	186	8.0 (4.7, 13.0)	1207	12.0 (10.0, 15.0)	994	16.7 (13.9, 19.7)
Total serum IgE > 140 kU/mL[c]	181	27.6 (20.6, 36.7)	984	19.4 (17.2, 22.2)	889	19.2 (16.6, 22.0)

Note. CI = confidence interval.
[a]The social class of a family was defined as the highest school grade completed by either the mother or the father of the study subject. Families with parental education of less than 10 grades were likewise assigned to social class I: 10 grades, class II; more than 10 grades, class III.
[b]Physician-diagnosed diseases as reported by parents.
[c]Perceived by parents.
[d]This value represents the 80th percentile of the entire population.

exposure. The effects of these potentially confounding variables were assessed in bivariate models that included each covariate separately. Additionally, covariates reaching significance levels of .2 or less in the bivariate analyses were selected for the final multivariate model.

Results

Lifetime prevalence rates for most of the major diseases and characteristic symptoms showed by the parents were highest in children from families with the highest level of parental education (social class III) and lowest in children from families with the lowest level of parental education (social class I). In children from social class II, prevalence rates were commonly at an intermediate level (Table 2).

In the physical examination, atopic dermatitis (eczema) was diagnosed in 2.0% of the children classified as belonging to social class III and in 3.3% of the children in social class II, while none of the 193 examined children in social class I presented signs of atopic dermatitis. Other health parameters assessed in the physical examination did not reveal significant differences between the social categories. Height, weight, and state of the teeth were comparable in all 3 groups (data not shown).

Allergic sensitization, defined as a positive response of 3 mm diameter or at least 1 of the 12 tested allergens, was found in 20.6% of the children. The allergic sensitization rates were 28.1%, 20.6%, and 11.8% in the children from social classes III, II, and I, respectively. Allergic sensitization against allergen groups or individual allergens as well as skin reactions to 2 or more allergens showed a similar social gradient (Table 2).

To assess the influence of total serum IgE on the relationship between social class and allergic sensitization, the subjects were stratified into a group with IgE levels above or equal to 140 IU/mL (90th percentile of the entire population) and a group whose IgE levels were below this cutoff point. Although children with high total serum IgE levels presented higher rates of allergic sensitization than children in the low IgE group, the effect of social class on allergic sensitization was evident in both strata (Figure 1).

The distribution of health-relevant factors and lifestyle characteristics among subjects in the 3 social classes is presented in Table 3. These factors were assessed to check for living conditions that may explain the observed differences in health outcomes between the social classes.

The influence of potentially confounding covariables on the association between allergic sensitization (dependent variable) and social class (explanatory independent variable) and the existence of interaction were analyzed for 1906 study subjects with complete data for the selected variables. Children in social classes II and III were compared with participants in social class I, which served as the reference group.

The unadjusted odds ratio for allergic sensitization comparing social class III with class I was 2.1, while the odds ratio comparing class II with class I was 2.0 (Table 4). Adjustment for individual covariables did not substantially influence the association between allergic sensitization and social class. Bivariate models containing covariables reaching a significance level of .1 are presented in Table 4. Interaction, assessed by including the each of the 9 variable models mentioned above the association variable between social class and the covariable under scrutiny, was not proved in any of the bivariate models. The final multivariate model included all variables selected a priori on theoretical grounds (sex, parental atopy, season when skin-prick test was performed, smoking exposure, and total serum IgE) and 2 other variables, study region and method of home housing, reached a significance level of .02 or less and met the criteria required for inclusion in the final multivariate model.

Total serum IgE (converted to the natural logarithm and treated as a continuous variable) had a considerable influence on the association between social class and allergic sensitization. Inclusion of the IgE variable in

TABLE 2—Lifetime Prevalence of Atopic Diseases and Characteristic Symptoms as Reported by Parents and Point Prevalence of Physician-Diagnosed Atopic Dermatitis, Allergic Sensitization, and Elevated Total Serum IgE in Children Aged 5 to 14 Years (n = 2402), by Social Class: Sachsen-Anhalt, Germany, 1992–1993

	Social class[a]					
	I		II		III	
	No.	Prevalence, % (95% CI)	No.	Prevalence, % (95% CI)	No.	Prevalence, % (95% CI)
Atopic disease or symptom (questionnaire information)						
Asthma or wheezy bronchitis[b]	191	4.7 (2.2, 8.8)	1173	7.8 (6.3, 9.4)	1021	10.8 (8.9, 12.8)
Wheezing (more than once)[c]	173	8.7 (4.9, 13.9)	1117	20.9 (18.5, 23.4)	961	22.8 (20.4, 25.6)
Eczema[b]	190	4.2 (1.8, 8.1)	1180	9.6 (8.0, 11.4)	1024	11.8 (9.9, 14.0)
Eczematous rashes[c]	189	8.5 (4.9, 13.4)	1170	14.1 (12.2, 16.2)	1021	16.5 (14.2, 18.9)
Allergic rhinitis[b]	190	1.1 (0.1, 3.8)	1180	4.1 (3.0, 5.4)	1024	5.0 (3.7, 6.5)
Allergic reaction[b]	190	10.5 (6.5, 15.8)	1170	12.6 (10.8, 14.7)	1023	15.3 (13.2, 17.7)
Atopic dermatitis (dermatologic examination)	163	0.0 (0.0, 2.2)	1053	3.3 (2.3, 4.6)	909	2.0 (1.2, 3.1)
Allergic sensitization (skin-prick test)						
At least 1 of 12 tested allergens	165	11.5 (7.1, 17.4)	1057	20.6 (18.2, 23.2)	904	22.1 (19.5, 25.0)
Mite allergens only (der p1 or der f1)	165	3.0 (1.0, 6.9)	1057	7.7 (6.1, 9.4)	904	8.3 (6.6, 10.3)
Fungi allergens only (aspergillus, alternaria)	165	5.5 (2.5, 10.1)	1057	5.2 (3.9, 6.7)	904	7.5 (5.9, 9.4)
Pollen (local grasses, hazel, birch, plantain, mugwort)	165	7.3 (3.8, 12.4)	1057	14.3 (12.2, 16.5)	904	15.5 (13.2, 18.0)
Cat epithelium	165	0.0 (0.0, 2.2)	1057	3.1 (2.2, 4.4)	904	3.7 (2.5, 5.1)
At least 2 of 12 tested allergens	165	8.5 (4.7, 13.8)	1057	12.9 (10.9, 15.0)	904	15.2 (12.9, 17.7)
Total serum IgE ≥ 140 IU/mL[d]	151	27.8 (20.8, 35.7)	964	19.6 (17.2, 22.2)	859	19.2 (16.6, 22.0)

Note. CI = confidence interval.

[a]The social level of a family was defined by the highest school grade completed by either the mother or the father of the study subject. Families with parental education of less than 10 grades were defined as belonging to social class I; 10 grades, class II; more than 10 grades, class III.

[b]Physician-diagnosed diseases as reported by parents.

[c]Perceived by parents.

[d]This value represented the 80th percentile of the entire population.

For Social Class I, Asthma or Wheezy bronchitis

$n = 191 \quad p = 0.047$ or 4.7%

1. *The hypothesis:* $H_0: P = 0.10$ vs $H_1: P \neq 0.10$

2. *The assumptions:* Independence

3. *The α-level :* $\alpha = 0.05$

4. *The test statistic:* $$z = \frac{p - 0.10}{\sqrt{\frac{p(1-p)}{n}}}$$

5. *The rejection region:* Reject $H_0: P = 0.10$, if z is not between ± 1.96

6. *The result:*

$$z = \frac{0.047 - 0.100}{\sqrt{\frac{0.047(1-0.047)}{191}}}$$

$$z = \frac{0.053}{\sqrt{\frac{0.047(0.953)}{191}}}$$

$$z = \frac{0.053}{\sqrt{\frac{0.0448}{191}}} = \frac{0.053}{\sqrt{0.0002}} = \frac{0.053}{0.0153}$$

$$z = 3.46$$

7. *The conclusion:* Reject H_0: P = 0.10 since z is not between ± 1.96

Confidence Interval, with $\alpha = 0.05$

$n = 191,\ p = 0.047,\quad p(1-p) = 0.047(0.953) = 0.0448$

$$C\left[p - 1.96\sqrt{\frac{p(1-p)}{n}} < P < p + 1.96\sqrt{\frac{p(1-p)}{n}}\right] = 0.95$$

$$C\left[0.047 - 1.96\sqrt{\frac{0.0448}{191}} < P < 0.047 + 1.96\sqrt{\frac{0.0448}{191}}\right] = 0.95$$

$$C\left[0.047 - 1.96(0.0153) < P < 0.047 + 1.96(0.0153)\right] = 0.95$$

$$C\left[0.047 - 0.03 < P < 0.047 + 0.03\right] = 0.95$$

$$C\left[0.017 < P < 0.077\right] = 0.95 \quad \text{or}$$

$$C\left[1.7\% < P < 7.7\%\right] = 0.95 \quad \text{vs} \quad C\left[2.2\% < P < 8.8\%\right] = 0.95$$