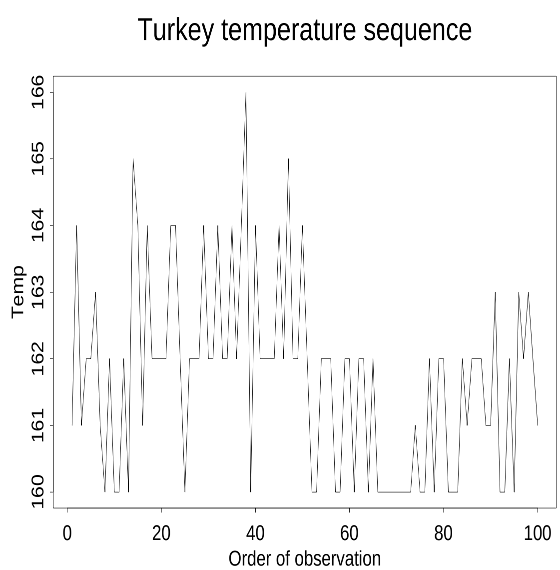
Analysis of turkey temperature data
and sample size recommendations
for Martin Bembers, Carolina Turkeys

Analyst: Jason A. Osborne, Dept. of Statistics, NCSU

January, 2002

Temperature measurements were taken on a sequence of $n = 100$ turkeys. The temperatures are plotted against the order in which the turkeys were measured in the figure below. This is NOT strictly the order in which the turkeys were taken out of the oven. Turkeys were taken out 10 or 11 at a time and measured. There is weak evidence ($p = 0.04$) of a weak lag 1 autocorrelation. That is, there appears to be a weak correlation $r \approx 0.2$ between the temperatures and the temperatures at the times immediately before them and this autocorrelation coefficient is significantly different from 0.



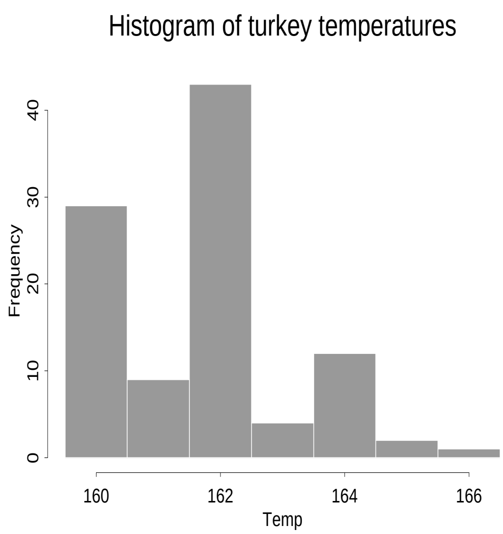
So, the $n = 100$ observations were treated as an effectively random sample in the exploration of the distribution of the temperature measurements. A histogram appears in the next figure. Note the high frequencies at $160, 162, 164^oF$. One possible explanation for this trimodal distribution is the precision of the thermometer. Rather than try to fit a trimodal distribution to the data, the measurements were binned into categories of width 2^oF:

Bin	$(159.5, 160.5)$	$(160.6, 162.5)$	$(162.6, 164.5)$	$(164.6, \infty)$
	160	$161, 162$	$163, 164$	$165, 166$
Observed Frequency	29	52	16	3
Expected Frequency	23.6	51.7	18.5	6.2

The expected values are estimated by fitting a gamma distribution to the shifted data. It is assumed that the minimum possible temperature in the oven from which these measurements were taken is $159.5^o F$, and that the distances above this minimum have a gamma distribution with unknown scale β and shape α parameters. The density function for this distribution is given below:

$$f(y; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}(y^{\alpha-1}e^{-\beta y} \text{ for } y > 0.$$

The maximum likelihood estimates of these parameters are $\hat{\alpha} = 1.94(0.22)$ and $\hat{\beta} = 0.88(0.09)$, with parametric bootstrap standard errors in parentheses. A chi-squared goodness-of-fit statistic of $\chi^2 = 2.27$ was calculated on $4 - 1 - 2 = 1$ degree of freedom for the table and did not show any lack-of-fit ($p > 0.1$).



Suppose now that the minimum temperature drops by $1^o F$ from 159.5 to 158.5, thus producing undercooked turkey breasts. How likely is it for a single turkey to be undercooked? Let the probability of such an event be denoted by π. The chance that at least one out of n turkeys will be measured below 160, assuming independence, is given by

$$\Pr(\text{detection}) = 1 - (1 - \pi)^n$$

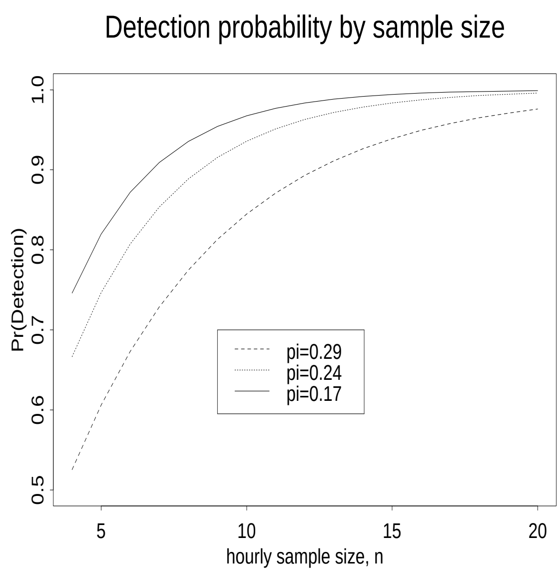
If X denotes a temperature randomly sampled from this out-of-control process, the MLE of $\pi(\alpha, \beta)$ is $\pi(\hat{\alpha}, \hat{\beta})$:

$$\hat{\pi} \quad = \quad \Pr(X < 159.5; \hat{\alpha}, \hat{\beta})$$

$$
\begin{aligned}
&= \Pr(y < 1.0; \hat{\alpha}, \hat{\beta}) \\
&= \int_0^1 \frac{\hat{\beta}^{\hat{\alpha}}}{\Gamma(\hat{\alpha})} y^{\hat{\alpha}-1} e^{-\hat{\beta} y} dy \\
&= 0.24
\end{aligned}
$$

(Note that another point estimate of this probability is just the observed proportion in the sample with temperature 160 or $\tilde{\pi} = 29/100 = 0.29$.) Estimated detection probabilities are given for various n in the following table

	$\hat{\pi} = 0.24$	$\tilde{\pi} = 0.29$	$\hat{\pi}_L = 0.17$
n	Pr(detection; $\hat{\pi}$)	Pr(detection; $\tilde{\pi}$)	Pr(detection; $\hat{\pi}_L$)
4	0.67	0.75	0.54
10	0.94	0.97	0.85
15	0.98	0.99	0.94
20	1.00	1.00	0.98



Note that the smaller the probability of detection for an individual turkey π, the smaller the chance to find at least one undercooked turkey. A conservative thing to do then is to find the sample size required for the smallest plausible value of π. A 95% confidence interval for π based on a parametric bootstrap is given by $(0.2, 0.3)$. An additional line is provided on the plot and in the table for the value $\pi = 0.2$. From these it can be seen that hourly samples of size $n = 20$ should ensure probability of detection at least 0.98. This *ad hoc* analysis suggests that $n = 4$ turkeys may be too few to detect problems when the minimum temperature falls from 159.5 to 158.5. Larger drops than $1^o F$ will not require such a large sample size.