Handout 2 — Fall 2010 — ACMS 20340

Why s^2 is an unbiased estimator for σ

There are two estimators, $\hat{\sigma}^2$ and s^2, for the population variance. Both work, and for n large enough, they are very close. The main difference is one is a biased estimator and the other isn't. This handout is for your own enjoyment, and is not necessary for the class.

Let our population have mean μ and standard deviation σ (and any distribution!). Suppose we have the following sample of size n:

$$\{x_1, x_2, \ldots, x_n\}$$

with each element of the sample labeled as x_i for some i. We already know how to find the sample average

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

We know $\bar{x}$ is a random variable which is normally distributed with mean μ and standard deviation $\sigma/\sqrt{n}$. As mentioned in class, we now need a way to estimate σ. We would like to do it with the estimator $\hat{\sigma}_\mu^2$ as follows

$$\hat{\sigma}_\mu^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

except we don't know μ! Since we are estimating μ with $\bar{x}$, so lets use that, giving the estimator $\hat{\sigma}^2$:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

And then we have the estimator s^2, which alters the fraction in front of $\hat{\sigma}^2$.

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Because $\hat{\sigma}^2$ is a random variable (its value depends on our random choice of sample), we can ask what its distribution, mean, and standard deviation are. These are not easy to answer. For one reason, the Central

Limit Theorem doesn't apply directly since we are *squaring* and adding instead of simply adding the samples together. However, some calculation (given below) does give the expected value

$$E(\hat{\sigma}^2) = \frac{n-1}{n}\sigma^2$$

The expected value of $\hat{\sigma}^2$ is not the variance σ^2 of our population! This means $\hat{\sigma}^2$ is biased estimator of σ^2. However, it is an unbiased estimator of $\frac{n-1}{n}\sigma^2$. Fine. If we multiply what our estimator says by $\frac{n}{n-1}$ then we should get an unbiased estimate of σ^2. Lets check:

$$E(\frac{n}{n-1}\hat{\sigma}^2) = \frac{n}{n-1}E(\hat{\sigma}^2) = \frac{n}{n-1}\frac{n-1}{n}\sigma^2 = \sigma^2$$

Perfect! So what exactly is $\frac{n}{n-1}\hat{\sigma}^2$? It is precisely the statistic s^2.

$$\begin{aligned} \frac{n}{n-1}\hat{\sigma}^2 &= \frac{n}{n-1}\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \\ &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \\ &= s^2 \end{aligned}$$

Remarks:

1. $\hat{\sigma}^2$ will give too small of an estimate for σ^2 on average (albeit, not too much smaller). Why is this? It is because we are using the sample mean $\bar{x}$ instead of μ in $\hat{\sigma}^2$. The estimator $\hat{\sigma}^2_\mu$ which uses the actual population mean instead of the sample mean has expected value σ^2, making it unbiased.

2. You probably noticed that the work above is with variance and not standard deviations. Why did we not work with the standard deviations directly? That is, doesn't $s = \sqrt{s^2}$ estimate σ? This is where things get complicated. Yes, we will use s to estimate σ. However, s is a biased estimator of σ! Good grief. Why? The square root function is to blame, because it doesn't behave as nice as dividing or multiplying. Using slightly more theory than we have covered in class will show
$$E(s) = E(\sqrt{s^2}) \leq \sqrt{E(s^2)} = \sigma$$
Thus using s to estimate σ (which we will do shamelessly and without reservations) underestimates the true value.

3. I promised a working out of $E(\hat{\sigma}^2)$. It is technical, and you certainly *do not* need to know it or understand it for class. We will need two properties of expected value, and one for variance

   (a) Since $Var(X) = E(X^2) - E(X)^2$ for any random variable X, we can rearrange to get $E(X^2) = Var(X) + E(X)^2$.

   (b) Expected values pass through multiplication and addition. That is, $E(aX + bY + c) = aE(X) + bE(Y) + c$, for any numbers a, b, c, and any random variables X and Y. We didn't talk about this in class.

   (c) Addition also passes through Variance, provided the random variables are independent. So $Var(X+Y) = Var(X) + Var(Y)$, if X and Y are independent random variables.

   Okay. Here goes:

$$
\begin{aligned}
E(\hat{\sigma}^2) &= E\left(\frac{1}{n}\sum(x_i - \bar{x})^2\right) \\
&= \frac{1}{n}E\left(\sum x_i^2 - n\bar{x}^2\right) \\
&= \frac{1}{n}\left[\sum E(x_i^2) - \frac{n}{n^2}E\left(\left(\sum x_i\right)^2\right)\right]
\end{aligned}
$$

   Now lets use the identity above to remove the $E(X^2)$'s.

$$
\begin{aligned}
&= \frac{1}{n}\left[\sum(\sigma^2 + \mu^2) - \frac{1}{n}\left(Var\left(\sum x_i\right) + E\left(\sum x_i\right)^2\right)\right] \\
&= \frac{1}{n}\left[n\sigma^2 + n\mu^2 - \frac{1}{n}\left(n\sigma^2 + (n\mu)^2\right)\right]
\end{aligned}
$$

   From now on it is just algebra

$$
\begin{aligned}
&= \frac{1}{n}\left[n\sigma^2 + n\mu^2 - \sigma^2 + n\mu^2\right] \\
&= \frac{1}{n}\left[(n-1)\sigma^2\right] \\
&= \frac{n-1}{n}\sigma^2
\end{aligned}
$$

   Exercise: use similar steps as above to show $E(\hat{\sigma}_\mu^2) = \sigma^2$.