GKC Aleutian sample size

Katie Palof, ADF&G

June 2, 2016

CPUE by string

Catch per unit effort (CPUE) was summarized by string, which includes 5 pots that were sampled in each string. CPUE is summarized in crab caught per hour of soak time. For reference, crab per day and SD of crab per day is also reported. CPUE by string, for total crab and for legal crab, shows the variability between the pots within a string. A representative sample of 18 strings is shown here.

Table 1: CPUE by string of total crab

string	CPUEbar	Csd	Cse	CV	crab_day	crab_day_sd	CV_base_rem
1	0.0550	0.0254	0.0114	46.18	1.3200	0.6096	33.84
2	0.1442	0.0660	0.0295	45.77	3.4608	1.5840	33.43
3	0.1937	0.0239	0.0107	12.34	4.6488	0.5736	0.00
4	0.1344	0.0436	0.0195	32.44	3.2256	1.0464	20.10
5	0.3062	0.0765	0.0342	24.98	7.3488	1.8360	12.64
6	0.1825	0.0684	0.0306	37.48	4.3800	1.6416	25.14
7	0.0779	0.0207	0.0093	26.57	1.8696	0.4968	14.23
8	0.2910	0.0748	0.0335	25.70	6.9840	1.7952	13.36
9	0.3322	0.0424	0.0190	12.76	7.9728	1.0176	0.42
10	0.1498	0.0595	0.0266	39.72	3.5952	1.4280	27.38
11	0.1077	0.0654	0.0292	60.72	2.5848	1.5696	48.38
12	0.1732	0.0493	0.0220	28.46	4.1568	1.1832	16.12
13	0.1660	0.1112	0.0497	66.99	3.9840	2.6688	54.65
14	0.0907	0.0139	0.0062	15.33	2.1768	0.3336	2.99
15	0.0929	0.0258	0.0115	27.77	2.2296	0.6192	15.43
16	0.1793	0.1419	0.0635	79.14	4.3032	3.4056	66.80
17	0.0949	0.0336	0.0150	35.41	2.2776	0.8064	23.07
18	0.0983	0.0164	0.0073	16.68	2.3592	0.3936	4.34

Key: **CPUEbar** is the mean CPUE for each string, **Csd** and **Cse** represent the standard deviation and standard error by string, **crab_day** & **crab_day_sd** are the mean and sd transformed to crab per day. **CV_base_rem** is the CV by string minus the minimum CV for all strings. This removes the underlying base variance and better reflects the true variability between pots in a string.

Table 2: CPUE by string of legal crab

string	CPUEbar	Csd	Cse	CV	crab_day	crab_day_sd	CV_base_rem
1	0.0530	0.0239	0.0107	45.09	1.2720	0.5736	29.28
2	0.0924	0.0319	0.0143	34.52	2.2176	0.7656	18.71
3	0.1234	0.0245	0.0110	19.85	2.9616	0.5880	4.04
4	0.0722	0.0254	0.0114	35.18	1.7328	0.6096	19.37
5	0.0982	0.0303	0.0136	30.86	2.3568	0.7272	15.05
6	0.1033	0.0452	0.0202	43.76	2.4792	1.0848	27.95
7	0.0574	0.0279	0.0125	48.61	1.3776	0.6696	32.80
8	0.2227	0.0358	0.0160	16.08	5.3448	0.8592	0.27
9	0.1066	0.0689	0.0308	64.63	2.5584	1.6536	48.82
10	0.0791	0.0333	0.0149	42.10	1.8984	0.7992	26.29
11	0.0654	0.0331	0.0148	50.61	1.5696	0.7944	34.80
12	0.0863	0.0236	0.0106	27.35	2.0712	0.5664	11.54
13	0.0778	0.0593	0.0265	76.22	1.8672	1.4232	60.41
14	0.0392	0.0166	0.0074	42.35	0.9408	0.3984	26.54
15	0.0732	0.0233	0.0104	31.83	1.7568	0.5592	16.02
16	0.0854	0.0487	0.0218	57.03	2.0496	1.1688	41.22
17	0.0945	0.0332	0.0148	35.13	2.2680	0.7968	19.32
18	0.0955	0.0151	0.0068	15.81	2.2920	0.3624	0.00

Key: **CPUEbar** is the mean CPUE for each string, **Csd** and **Cse** represent the standard deviation and standard error by string, **crab_day** & **crab_day_sd** are the mean and sd transformed to crab per day. **CV_base_rem** is the CV by string minus the minimum CV for all strings. This removes the underlying base variance and better reflects the true variability between pots in a string.

CPUE by region

For total crab per pot

```
## Source: local data frame [1 x 6]
##
##   CPUE_region  sd_region        Cse       CV crab_day crab_day_sd
##         (dbl)      (dbl)      (dbl)    (dbl)    (dbl)       (dbl)
## 1   0.1594389 0.08017144 0.01889659 50.28349 3.826533    1.924115
```

For legal crab per pot

```
## Source: local data frame [1 x 6]
##
##   CPUE_region  sd_region         Cse       CV crab_day crab_day_sd
##         (dbl)      (dbl)       (dbl)    (dbl)    (dbl)       (dbl)
## 1  0.09031111 0.03897073 0.009185488 43.15164 2.167467   0.9352974
```

Power / Sample size

Analysis were done to determine the effect size of the current sample size of 66 pots to detect a change in regional CPUE from year to year. The effect size will determine how large of a change in CPUE the current

sample size can detect from year to year. Additionally, the effect size and power of the current sample size of 5 pots per string was reviewed. The effect size to detect a difference between strings was determined, and the possibility of adding more samples per string was examined.

regional sample size (number of strings)

The power to determine a difference in regional CPUE from year to year. This analysis relies on the standard deviation of the mean (here CPUE) for the region. Initially other parameters (Power = 0.85 and sig.level = 0.05) were held constant to determine the effect size of the current sample size.

Effect size for a sample size of 66 strings (assuming 85% power and a significance level of 0.05)

```
##
##      Two-sample t test power calculation
##
##               n = 66
##               d = 0.5255064
##       sig.level = 0.05
##           power = 0.85
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

For total crab CPUE (assuming a regional SD for total crab of 0.0802), this effect size equates to an ability to detect a difference in CPUE that is larger than the below value in crabs per day.

```
## [1] 1.011495
```

For legal crab CPUE(assuming a regional SD for LEGAL CPUE of 0.0390), this effect size equates to an ability to detect a difference in CPUE that is larger than the below value in crabs per day.

```
## [1] 0.491874
```

The power of the current sampling regime depends on what the desired effect size (or difference to detect between CPUEs) is. The current sample size of 66 strings has the ability to detect a change of 0.49 LEGAL crab per day or greater with a power of 85% for a significance level of 0.05.

pots per string sample size

Using the average standard deviation observed within strings, the effect size can be determined at 85% power to detect a significance level of 0.05. This is the power to detect a difference (effect size) in CPUE between strings.

```
##
##      Two-sample t test power calculation
##
##               n = 5
##               d = 2.167515
##       sig.level = 0.05
##           power = 0.85
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

For legal crab CPUE(assuming an average SD within pots of 0.0333), this effect size equates to an ability to detect a difference in CPUE that is larger than the below value in crabs per day between pots.

```
## [1] 1.734012
```

This value is fairly large (considering CPUE of crab per day for LEGAL crab ranges from 0.94 t0 5.34), so additional power analyses were performed to determine the sample size necessary to get a smaller effect size - closer to 0.5 crab per day using the same power and significance level.

```
##
##      Two-sample t test power calculation
##
##               n = 46.94997
##               d = 0.625
##       sig.level = 0.05
##           power = 0.85
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

This produces a larger sample size than possible since there are only 35 pots per string. Variability within a string is to high to efficiently detect this small of a change in CPUE between strings.

Assuming a sample size of 5 pots per string, and an effect size of 0.5 crab per day, the power to detect this difference between string is:

```
##
##      Two-sample t test power calculation
##
##               n = 5
##               d = 0.625
##       sig.level = 0.05
##           power = 0.1410202
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

The current sample size of 5 pots per string has fairly low power (~14%) to detect a difference of 0.5 crab per day between strings at a significance level of 0.05.

summay of pots per string

In general, the number of pots sampled per string could be increased, however the variability within each string is so high that there are diminishing returns for increasing the number of pots sampled. It is recommended that the number of pots sampled per string remain at 5 since this appears sufficient enough to look at CPUE at a regional level due to the larger number of strings sampled (n=66).

Additionally, when examining the CV for each string it appears that there is a baseline level of variability (~ a CV of 16 for LEGAL crab) present in each string. When this baseline level is removed (CV_base_rem) a more realistic view of the variability within strings can be observed. The average CV (for LEGAL crab) with the baseline removed hovers around 24, which is an acceptable level of variability for this sampling method. Increasing the number of pots sampled would NOT decrease this variability substantial and therefore is NOT recommended.