Risk-based sampling: Counting beans

Christina Devorshak
USDA APHIS PPQ ST

International Symposium for
RISK-BASED SAMPLING
June 26-30, 2017 • Baltimore, Maryland

Task: how to explain how RBS works to a non-technical audience?

- Exercise created to be used in workshops
- Easily implemented
- Clearly demonstrates RBS principles
- Repeatable / consistent results
- Tangible activity aids comprehension and "believability"

Bean counting

Exercise

- You have been provided "consignments" of 100, 500, 1,000, 2,000, and 5,000 beans with a 10% infestation rate of black beans
- **Read your instruction sheets!**
- Blindly sample beans from each bag
- **Do not replace beans** until each sample is completed
- Record results and replace all beans removed from bag
- Shake the bag
- **Repeat three times**
- Average results

Bean counting

Calculating results

- Calculate the sampling rate required for detection (total number of beans sampled divided by the size of the consignment)

- Express results as a percentage

- Share results with other participants

Hypergeometric tables

Optimum Sample Sizes Given: Lot Size (N) Acceptable Risk Levels (r) Confidence Levels (p) N=100						
Infestation rate (r) (acceptable risk level)	**Confidence level (p)**					
	0.8	**0.85**	**0.9**	**0.95**	**0.99**	**0.999**
0.0100	80	85	90	95	99	100
0.0500	27	31	37	45	59	74
0.1000	15	17	20	25	36	48

= risk based sampling rate

≈ 2% sampling rate

Optimum Sample Sizes Given:
Lot Size (N)
Acceptable Risk Levels (r)
Confidence Levels (p)

N =   100

Acceptable Risk Level (r)	Confidence Level (p)					
	0.8	0.85	0.9	0.95	0.99	0.999
0.0100	80	85	90	95	99	100
0.0200	56	61	69	78	90	97
0.0300	42	47	54	63	78	90
0.0400	33	38	44	52	68	81
0.0500	27	31	37	45	59	74
0.0600	23	27	32	39	53	67
0.0700	20	24	28	34	47	61
0.0800	18	21	25	31	43	56
0.0900	16	19	22	28	39	52
0.1000	15	17	20	25	36	48
0.1100	13	16	18	23	33	45
0.1200	12	14	17	21	31	42
0.1300	11	13	16	20	29	39
0.1400	11	12	15	19	27	37
0.1500	10	12	14	17	25	35
0.1600	9	11	13	16	24	33
0.1700	9	10	12	15	22	31
0.1800	8	10	11	15	21	30
0.1900	8	9	11	14	20	28
0.2000	7	9	10	13	19	27

☐ = risk based sampling rate

☐ ≈ 2% sampling rate

Optimum Sample Sizes Given:
            Lot Size (N)
            Acceptable Risk Levels (r)
            Confidence Levels (p)

                    N =   500

Acceptable Risk Level (r)	Confidence Level (p)					
	0.8	0.85	0.9	0.95	0.99	0.999
0.0100	138	158	184	225	300	373
0.0200	74	86	102	129	183	248
0.0300	51	59	71	90	131	182
0.0400	38	45	54	69	101	144
0.0500	31	36	43	56	83	118
0.0600	26	30	36	47	70	100
0.0700	22	26	31	40	60	87
0.0800	19	23	27	35	53	77
0.0900	17	20	24	31	47	69
0.1000	16	18	22	28	42	62
0.1100	14	17	20	26	38	56
0.1200	13	15	18	23	35	52
0.1300	12	14	17	22	33	48
0.1400	11	13	16	20	30	44
0.1500	10	12	14	19	28	41
0.1600	10	11	14	17	26	39
0.1700	9	11	13	16	25	36
0.1800	9	10	12	15	23	34
0.1900	8	9	11	15	22	32
0.2000	8	9	11	14	21	31

Acceptable Risk Levels (r)
Confidence Levels (p)

N =   1000

Acceptable Risk Level (r)	Confidence Level (p)					
	0.8	0.85	0.9	0.95	0.99	0.999
0.0001	1000	1000	1000	1000	1000	1000
0.0002	1000	1000	1000	1000	1000	1000
0.0003	996	999	1000	1000	1000	1000
0.0004	983	992	998	1000	1000	1000
0.0005	961	978	991	998	1000	1000
0.0006	932	958	979	994	1000	1000
0.0007	900	934	963	987	999	1000
0.0008	867	907	944	977	997	1000
0.0009	833	879	923	965	995	1000
0.0010	800	850	900	950	990	999
0.0020	553	613	684	777	900	968
0.0030	415	469	536	631	784	900
0.0040	331	378	438	527	683	821
0.0050	275	316	369	450	601	748
0.0060	235	271	318	393	535	683
0.0070	205	237	280	348	481	626
0.0080	182	211	250	312	437	577
0.0090	164	190	225	282	399	534
0.0100	148	173	205	258	368	497
0.0200	77	90	108	138	204	290
0.0300	52	61	73	94	141	203
0.0400	39	46	55	71	107	156
0.0500	31	37	44	57	86	126
0.0600	26	31	37	48	72	106
0.0700	22	26	32	41	62	91
0.0800	20	23	28	36	54	80
0.0900	17	20	25	32	48	71
0.1000	16	18	22	29	43	64
0.1100	14	17	20	26	39	58
0.1200	13	15	18	24	36	53
0.1300	12	14	17	22	33	49
0.1400	11	13	16	20	31	45
0.1500	10	12	15	19	28	42
0.1600	10	11	14	18	27	39
0.1700	9	11	13	16	25	37
0.1800	9	10	12	16	23	35
0.1900	8	9	11	15	22	33
0.2000	8	9	11	14	21	31

Lot Size (N)
Acceptable Risk Levels (r)
Confidence Levels (p)

N = 5000

Acceptable Risk Level (r)	Confidence Level (p)					
	0.8	0.85	0.9	0.95	0.99	0.999
0.0001	4801	4888	4951	4988	5000	5000
0.0002	4000	4250	4500	4750	4950	4995
0.0003	3290	3589	3923	4322	4768	4950
0.0004	2764	3064	3419	3882	4500	4842
0.0005	2374	2659	3010	3491	4207	4684
0.0006	2076	2343	2679	3158	3922	4500
0.0007	1843	2092	2410	2875	3658	4305
0.0008	1656	1888	2188	2635	3418	4110
0.0009	1503	1720	2002	2430	3202	3922
0.0010	1376	1579	1845	2253	3009	3743
0.0020	743	864	1028	1294	1844	2492
0.0030	508	594	711	904	1320	1843
0.0040	386	452	543	695	1027	1458
0.0050	311	365	439	564	840	1205
0.0060	261	306	369	474	710	1026
0.0070	224	263	318	409	615	893
0.0080	197	231	279	360	542	790
0.0090	175	206	249	321	485	709
0.0100	158	186	224	290	438	643
0.0200	80	94	113	147	223	331
0.0300	53	62	76	98	149	222
0.0400	40	47	57	73	112	167
0.0500	32	37	45	59	89	133
0.0600	26	31	38	49	74	111
0.0700	23	27	32	42	64	95
0.0800	20	23	28	36	55	83
0.0900	18	21	25	32	49	73
0.1000	16	18	22	29	44	66
0.1100	14	17	20	26	40	59
0.1200	13	15	19	24	36	54
0.1300	12	14	17	22	34	50
0.1400	11	13	16	20	31	46
0.1500	10	12	15	19	29	43
0.1600	10	11	14	18	27	40
0.1700	9	11	13	17	25	38
0.1800	9	10	12	16	24	35
0.1900	8	10	11	15	22	33
0.2000	8	9	11	14	21	31

Sampling variables
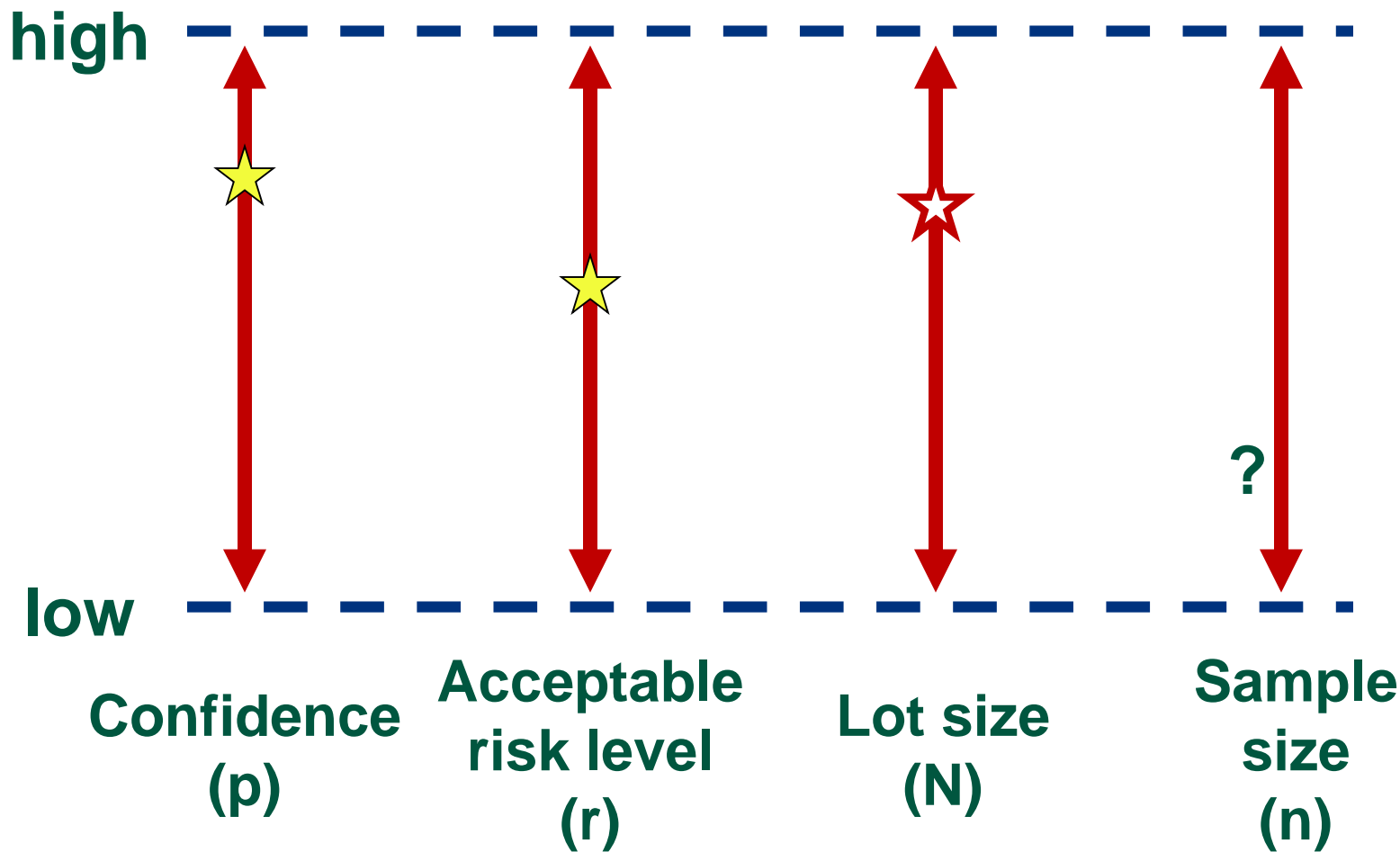
high ---- ---- ---- ---- ---- ---- ---- ---- ---- ----

low ---- ---- ---- ---- ---- ---- ---- ---- ---- ----

Confidence (p)

Acceptable risk level (r)

Lot size (N)
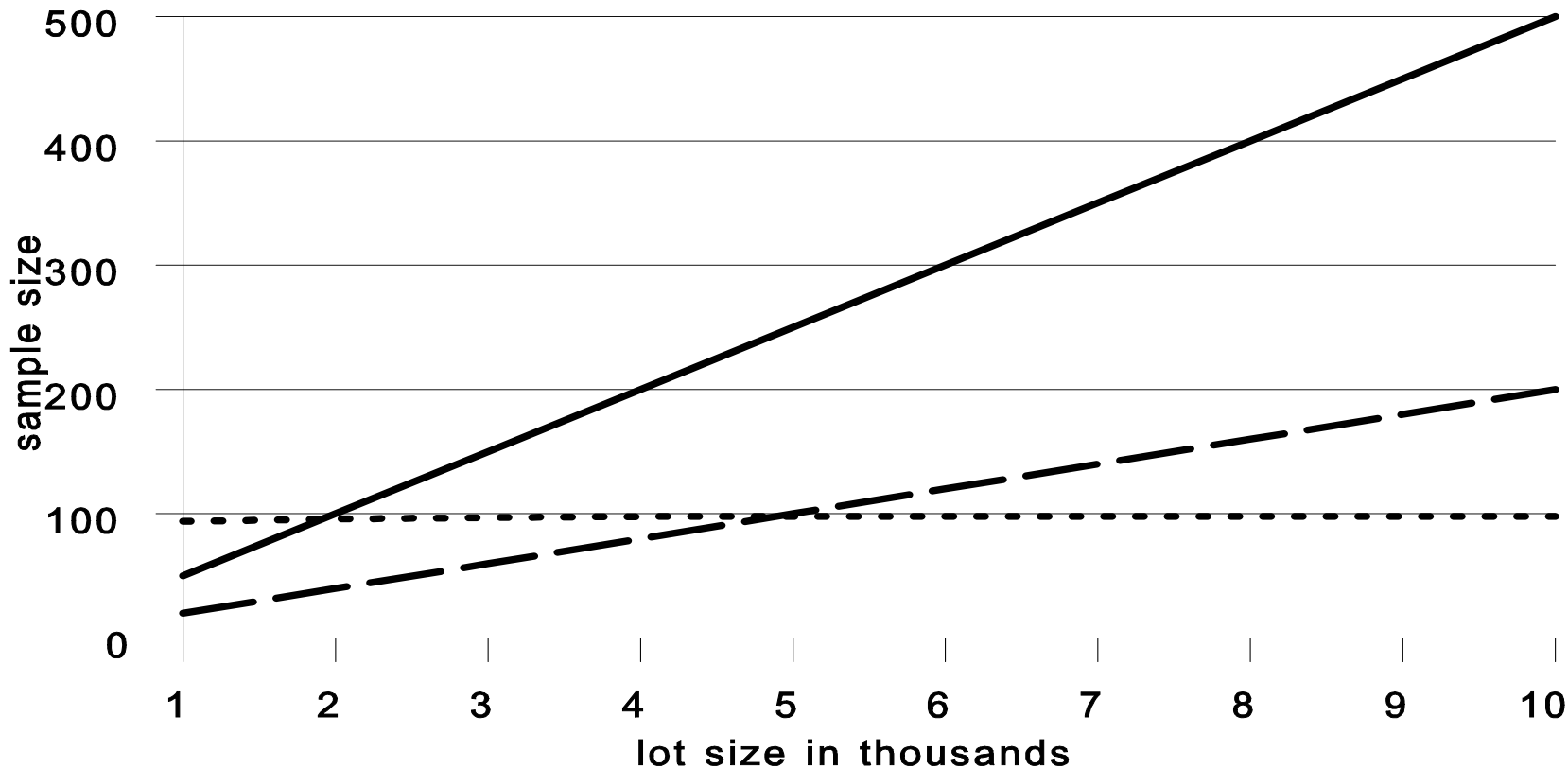
Sample size (n) ?
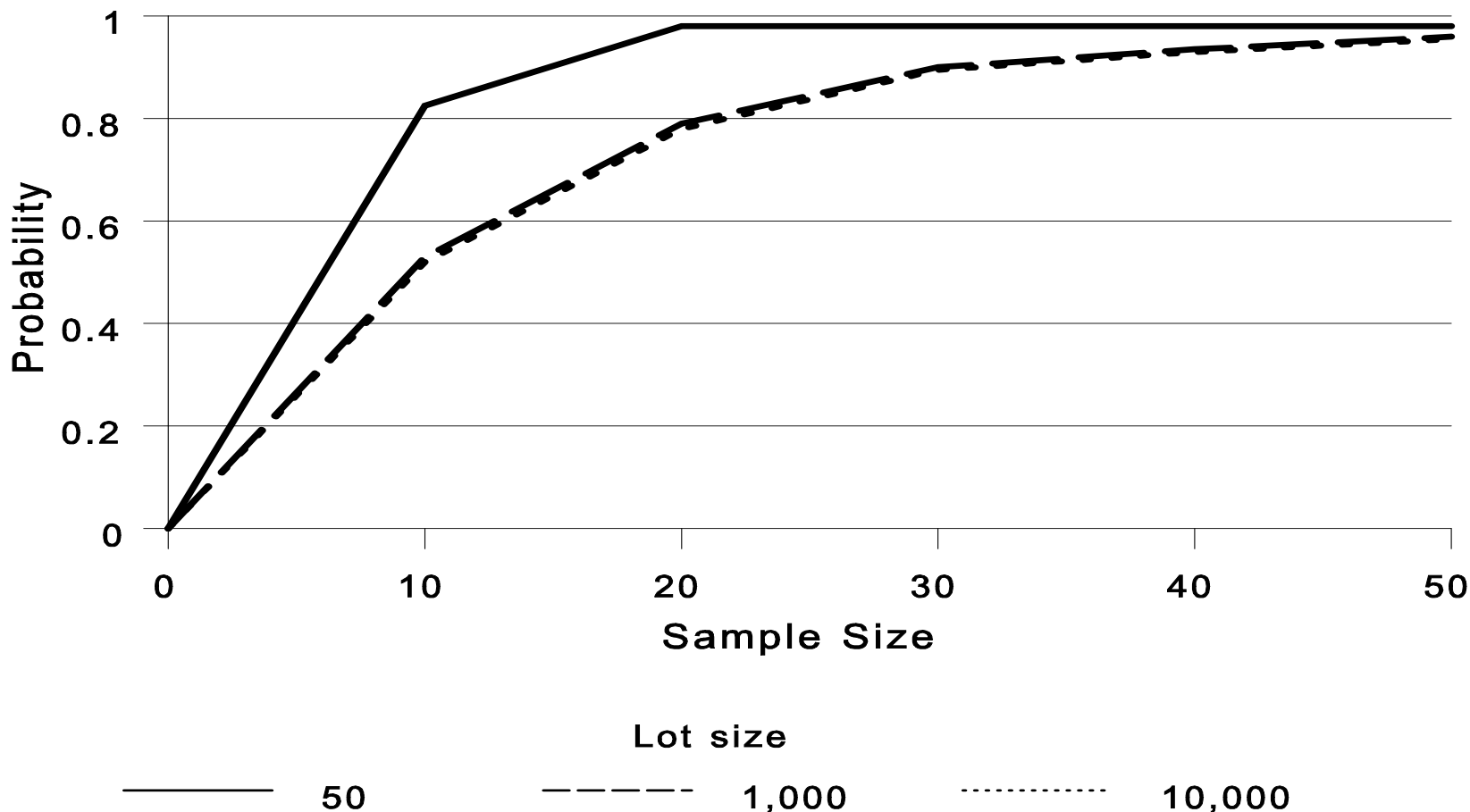
Probability of Detection

(based on infestation rate of 1/10)

Sample size observations

- The probability of detection is significantly different in the 0-30 sample range
- The probability of detection is not significantly different above 30 samples

Sample size observations

- The probability of detection for the small lot changes very little after 20 samples
- The probability of detection for the large lots changes very little after 50 samples