Практическое занятие 10: Оценка генеральной средней и генеральной доли по повторной и бесповторной выборке

Цель занятия: приобретение навыков оценки генеральной средней и генеральной доли по повторной и бесповторной выборке.

Бесповторная выборка – если случайно отбираемые объекты не возвращаются в генеральную совокупность.

Повторная выборка – если выбранный объект возвращается обратно (перед выбором следующего) в генеральную совокупность. Здесь один и тот же объект может быть выбран неоднократно.

Привычнее и понятнее *бесповторный отбор* (примеры):

- оценка качества партии помидоров в партии (повторный отбор ухудшит репрезентативность выборки);
- исследование успеваемости студентов ВУЗа;
- телефонный опрос *«Верите ли вы в Деда Мороза?»*.

Примеры повторного отбора:

- статистическое исследование прогулов в ВУЗе (один и тот же студент может попасть в выборку неоднократно, и было неправильно не учитывать его повторные прогулы;
- количество обращений в поликлинику (один тот же человек может обратиться несколько раз);
- многократное измерение некоторой величины. Теоретически генеральная совокупность бесконечна, и из неё мы «выбираем» несколько значений, которые могут повторяться, причём, не только теоретически, но и практически, по причине округления измерений.

Полагаем, что во всех нижеследующих задачах *генеральная совокупность* распределена нормально, либо её распределение близко к таковому. Этот факт может быть известен и / или подкреплён статистическими методами.

Оценка генеральной средней

Пусть из **нормально распределенной** генеральной совокупности объёма $\boldsymbol{N}$ проведена выборка объёма $\boldsymbol{n}$ и по её результатам найдена выборочная средняя $\overline{x}_в$ и исправленная выборочная дисперсия $\boldsymbol{s^2}$.

Тогда **доверительный интервал** для оценки *генеральной средней* $\overline{x}_в = a$ имеет вид: $\overline{x}_в - \Delta < \overline{x}_Г < \overline{x}_в + \Delta$, где **Δ** *(«дельта»)* – **точность оценки**, которую также называют ***предельной ошибкой*** *репрезентативности выборки*.

Точность оценки рассчитывается как произведение $\boldsymbol{\Delta = t \cdot \mu}$ – *коэффициента доверия* **t** на ***среднюю ошибку*** выборки $\boldsymbol{\mu}$ *(«мю»)*.

Для бесповторной выборки составляет $\mu = \sqrt{\frac{s^2}{n}\left(1 - \frac{n}{N}\right)}$,

а для повторной: $\mu = \sqrt{\frac{s^2}{n}}$.

В том случае, если изначально известна **генеральная дисперсия** $\boldsymbol{\sigma^2}$, то используют, конечно, её.

Если объём выборки n<30, то коэффициент доверия определяется **с помощью распределения Стьюдента**.

Если $\boldsymbol{n} \geq 30$, то чаще пользуются соотношением $\boldsymbol{2\Phi(t) = \gamma}$, где $\boldsymbol{\Phi(x)}$ – **функция Лапласа**, а $\boldsymbol{\gamma}$ – **доверительная вероятность**. Значение «гамма» показывает, с какой вероятностью построенный интервал $(\overline{x}_в - \Delta; \overline{x}_в + \Delta)$ накрывает истинное значение $\overline{x}_Г$.

Задание 1

С целью изучения урожайности подсолнечника в колхозах области проведено 5%-ное выборочное обследование 100 га посевов, отобранных в случайном порядке, в результате которого получены следующие данные:

Урожайность, ц/га	Посевная площадь, га
до 13	10
от 13 до 15	25
от 15 до 17	40
от 17 до 19	20
свыше 19	5
$\sum=$	100

С вероятностью 0,9973 определить *предельную ошибку* выборки и возможные границы, в которых ожидается средняя урожайность подсолнечника в области.

Методика выполнения

В условии не указан тип отбора, но исходя из логики исследования, положим, что он *бесповторный*. Поскольку выборка 5%-ная, то она составляет 1/20-ю часть генеральной совокупности, стало быть, общая посевная площадь области составляет: N=(100%/5%)n= 20*100=2000 га.

По условию, требуется найти *предельную ошибку* выборки $\boldsymbol{\Delta = t \cdot \mu}$, где $\boldsymbol{t}$ – коэффициент доверия, соответствующий доверительной вероятности γ=0,9973, и коль скоро, выборка *бесповторна* и **генеральная дисперсия** не неизвестна, то *средняя ошибка* рассчитывается по формуле $\mu = \sqrt{\frac{s^2}{n}\left(1 - \frac{n}{N}\right)}$.

Далее нужно найти интервал $(\overline{x}_в - \Delta; \overline{x}_в + \Delta)$, который с вероятностью 99,73% накроет *генеральную среднюю* $\overline{x}_Г$ урожайность подсолнечника по области.

Не известна ни сама **выборочная средняя**, ни **исправленная выборочная дисперсия**.

Можно сделать вывод, что предложен **интервальный вариационный ряд** с открытыми крайними интервалами. Поскольку длина частичного интервала составляет $\boldsymbol{h}$=2 га, то вопрос закрываем так: 11-13 и 19-21 га.

Находим середины $\boldsymbol{x_i}$ интервалов (переходим к **дискретному ряду**), произведения $\boldsymbol{x_i n_i}$, $\boldsymbol{x_i^2 n_i}$ и их суммы:

Интервалы	x_i	n_i	$x_i n_i$	$x_i^2 n_i$
11-13	12	10	120	1440
13-15	14	25	350	4900
15-17	16	40	640	10240
17-19	18	20	360	6480
19-21	20	5	100	2000
Σ=		100	1570	25060

$$\overline{x}_e = \frac{\sum_{i=1}^{k} x_i n_i}{n} = 1570/100 = \mathbf{15{,}7}$$

$$\overline{\sigma}_в = \frac{\sum_{i=1}^{k} x_i^2 n_i}{n} - (\overline{x}_в)^2 = \mathbf{4{,}11}$$

Исправленная дисперсия:

$$\varepsilon^2 = \frac{n}{n-1}\delta_в^2 = \frac{100}{99}4{,}11 = \mathbf{4{,}152}$$

И составляем доверительный интервал $(\overline{x}_в - \Delta; \overline{x}_в + \Delta)$ для оценки генеральной средней урожайности подсолнечника по области.

Вычислим предельную ошибку ***Δ=t•μ***.

Так как объём выборки ***n***>30, то коэффициент доверия ищем из соотношения ***2Φ(t)=γ***. Поскольку ***γ***=0,9973, то: Φ(t)=0,9973/2=0,49865.

По **таблице значений функции Лапласа** определяем, что этому значению функции соответствует аргумент ***t***=3.

Вычислим *среднюю ошибку* бесповторной выборки:

$$\mu = \sqrt{\frac{s^2}{n}\left(1-\frac{n}{N}\right)} = \sqrt{\frac{4{,}15}{100}\left(1-\frac{100}{2000}\right)} \approx 0{,}20\ ц/га.$$

Таким образом, предельная ошибка составляет ***Δ=t•μ***=3*0,20=0,6 ц/га, и искомый доверительный интервал:

$$(\overline{x}_в - \Delta < \overline{x}_Г < \overline{x}_в + \Delta)$$

$$(15{,}7 - 0{,}6 < \overline{x}_Г < 15{,}7 + 0{,}6)$$

$$15{,}1 < \overline{x}_Г < 16{,}3$$

– границы, в которых ожидается средняя урожайность подсолнечника в области с вероятностью ***γ***=0,9973.

Кстати, такое «странное» значение вероятности не случайно, дело в том, что оно соответствует **правилу «трёх сигм»**, т.е., практически достоверным является тот факт, что построенный интервал накроет истинное значение $\overline{x}_Г$ средней урожайности по области.

Ответ: ***Δ**=3*0,20=0,6* ц/га, $15{,}1 < \overline{x}_Г < 16{,}3$ (ц/га).

Теперь распишем интервал в развёрнутом виде:

$$\overline{x}_в - t\sqrt{\frac{s^2}{n}\left(1-\frac{n}{N}\right)} < \overline{x}_Г < \overline{x}_в + t\sqrt{\frac{s^2}{n}\left(1-\frac{n}{N}\right)}$$

и проанализируем дробь ***n/N***. Очевидно, что при увеличении *объёма выборки* ***n*** эта дробь будут увеличиваться до единицы, и, соответственно, разность ***1-n/N*** будет уменьшаться до нуля.

Таким образом, предельная ошибка $\Delta = t\sqrt{\frac{s^2}{n}\left(1-\frac{n}{N}\right)}$ уменьшается, и доверительный интервал становится меньше, что вполне логично – ведь чем больше выборка, тем точнее оценка.

В предельном случае, когда исследована вся генеральная совокупность ***n=N***, ошибка Δ становится нулевой и доверительный интервал *вырождается* в генеральную среднюю $\overline{x}_Г$.

Исходя из вышесказанного, можно рассмотреть две обратные задачи:

1) Предположим, что нужно уменьшить доверительный интервал, например, в два раза, т.е. споловинить предельную ошибку до Δ=0,3 ц/га (вместо 0,6). Но высокую доверительную вероятность и соответствующий коэффициент ***t***=3 нужно сохранить. Тогда ничего не остаётся, как увеличивать объём выборки. Из соотношения $\Delta = t\sqrt{\frac{s^2}{n}\left(1-\frac{n}{N}\right)}$ выведем формулу для нахождения этого объёма, для этого возведём обе части в квадрат:

$$\Delta^2 = t^2 s^2\left(1-\frac{n}{N}\right) \;\Rightarrow\; n = \frac{s^2 t^2 N}{s^2 t^2 + \Delta^2 N}$$

Таким образом, для того чтобы с доверительной вероятностью γ =0,9973 обеспечить точность Δ=0,3, следует организовать выборку объёмом: $n = \frac{s^2 t^2 N}{s^2 t^2 + \Delta^2 N} = \frac{4{,}15*3*3*2000}{4{,}15*3*3+0{,}3*0{,}3*2000} \approx 344$ гектара.

Округляем в большую сторону, что составляет 344/2000=17,2% генеральной совокупности. Таким образом, трудозатраты возрастут примерно в 3,5 раза.

Следует обратить **ОСОБОЕ внимание**, что найденный ранее интервал $15{,}1 < \overline{x}_Г < 16{,}3$ уменьшать в два раза **НЕЛЬЗЯ**: $\cancel{15{,}4 < \overline{x}_Г < 16}$.

Потому что в новой выборке почти наверняка будет получено **другое** значение $\overline{x}_в$ и интервал $(\overline{x}_в - \Delta; \overline{x}_в + \Delta)$ «сдвинется», да и точность Δ=0,3 будет выдержана лишь примерно (т.к. значение ***s²*** тоже изменится).

2) Теперь обратная ситуация – когда оценка Δ=0,6 устраивает, но нет возможности или времени проводить большую выборку. Поэтому и исследовали ***n***=20 гектаров из ***N***=2000. В этом случае пострадает доверительная вероятность, давайте выясним насколько:

$$\Delta = t\sqrt{\frac{s^2}{n}\left(1-\frac{n}{N}\right)} \quad 0{,}6 = t\sqrt{\frac{4{,}15}{20}\left(1-\frac{20}{2000}\right)} \Rightarrow t = \frac{0{,}6}{\sqrt{0{,}2075 * 0{,}99}} \approx 1{,}32$$

и по **таблице значений функции Лапласа** находим:

$\boldsymbol{\gamma = 2\Phi(t)} \approx 2\Phi(1{,}132)=2*0{,}4966=0{,}8132$.

Впрочем, это было очевидно – ведь такая малая выборка явно *не репрезентативна* (плохо представляет генеральную совокупность).

Варианты задания 1

№ вар.	Урожайность, ц/га					% обслед.	p
	до 13	от 13 до 15	от15 до 17	от 17 до 19	свыше 19		
	Посевная площадь, га						
1	36	48	26	27	28	5,00	0,978
2	44	17	48	24	50	9,00	0,983
3	11	28	17	9	7	6,00	0,920
4	41	10	44	15	46	6,00	0,824
5	20	37	21	24	7	8,00	0,929
6	40	29	9	28	36	9,00	0,898
7	48	33	8	16	45	8,00	0,949
8	13	24	41	32	33	7,00	0,943
9	43	26	22	11	26	7,00	0,886
10	15	32	48	28	7	7,00	0,901
11	13	21	11	9	29	8,00	0,813
12	10	49	32	10	35	8,00	0,908
13	49	50	41	6	7	8,00	0,906
14	9	13	48	17	26	9,00	0,980
15	43	43	50	27	12	5,00	0,902
16	37	49	26	6	20	6,00	0,942
17	49	46	12	37	38	6,00	0,842
18	35	40	25	27	40	5,00	0,832
19	31	34	43	38	29	8,00	0,946
20	22	13	41	36	25	7,00	0,980
21	27	34	29	36	7	9,00	0,836
22	50	39	46	22	10	3,00	0,937
23	14	43	43	33	39	3,00	0,823
24	31	10	8	39	5	8,00	0,868
25	7	44	50	28	42	7,00	0,855

Задание 2

По результатам 10%-ной бесповторной выборки объёма $\boldsymbol{n}$=50, найдены выборочная средняя $\overline{x}_{в}$=107,92 и дисперсия $\boldsymbol{D_в}$=84,51.

а) Найти пределы, за которые с доверительной вероятностью 0,954 не выйдет среднее значение генеральной совокупности.

б) Выборку примерно какого объёма нужно организовать, чтобы с той же доверительной вероятностью улучшить точность оценки в три раза?

Методика выполнения

вычислим исправленную выборочную дисперсию:

$$s^2 = \frac{n}{n-1} D_в = \frac{50}{49} 84{,}51 \approx 86{,}23$$

а) Вычислим предельную ошибку $\boldsymbol{\Delta = t \bullet \mu}$ выборки. Так как $\boldsymbol{n}$>30, то коэффициент доверия найдём из соотношения $\boldsymbol{\gamma = 2\Phi(t)}$.

По условию, $\boldsymbol{\gamma}$ =0,954, следовательно: $\boldsymbol{2\Phi(t)}$=0,954 $\Rightarrow$ $\boldsymbol{\Phi(t)}$= 0,477. По **таблице значений функции Лапласа** находим, что этому значению функции соответствует аргумент $\boldsymbol{t}$=2,

Поскольку выборка 10%-ная бесповторная, то объём генеральной совокупности равен: $\boldsymbol{N}$= (100%/10%)*$\boldsymbol{n}$= 500.

Вычислим среднюю ошибку выборки:

$$\mu = \sqrt{\frac{s^2}{n}\left(1 - \frac{n}{N}\right)} = \sqrt{\frac{86{,}23}{50}\left(1 - \frac{50}{500}\right)} \approx 1{,}246.$$

Таким образом, предельная ошибка:

$\boldsymbol{\Delta = t \bullet \mu}$ ≈2*1,246≈2,49 и искомый доверительный интервал:

$$(\overline{x}_в - \Delta < \overline{x}_Г < \overline{x}_в + \Delta)$$

$$(107{,}92 - 2{,}49 < \overline{x}_Г < 107{,}92 + 2{,}49)$$

$$105{,}43 < \overline{x}_Г < 110{,}41$$

– пределы, которые с доверительной вероятностью 0,954 накрывают среднее значение генеральной совокупности.

б) Улучшим точность оценки в три раза: $\boldsymbol{\Delta}$=2,49/3=0,83 и воспользуемся формулой:

$$n = \frac{s^2 t^2 N}{s^2 t^2 + (\Delta^*)^2 N} \approx \frac{86{,}23 * 2 * 2 * 500}{86{,}23 * 2 * 2 + 0{,}83^2 * 500} \approx 251$$

(округлять лучше до большего значения).

Таким образом, для того, чтобы с вероятностью 95,4% утверждать, что $\overline{x}_в$ отличается от $\overline{x}_Г$ менее чем на $\boldsymbol{\Delta}$=0,83, следует провести выборку объёмом примерно $\boldsymbol{n}$=251 (что составляет половину генеральной совокупности, и, конечно, нецелесообразно).

Ответ: а) $105,43 < \bar{x}_Г < 110,41$, б) $\boldsymbol{n}$=251.

Варианты задания 2

№ вар.	n	% выборки	$\bar{x}_в$	$D_в$	$P_{доверит.}$	№ вар.	n	% выбор	$\bar{x}_в$	$D_в$	$P_{доверит.}$
1	**48**	**9**	**106**	**73**	**0,9597**	**14**	**43**	**10**	**88**	**71**	**0,8808**
2	**38**	**9**	**95**	**76**	**0,8271**	**15**	**25**	**11**	**122**	**80**	**0,8178**
3	**35**	**10**	**115**	**74**	**0,9227**	**16**	**54**	**10**	**95**	**79**	**0,9046**
4	**46**	**11**	**95**	**79**	**0,8971**	**17**	**27**	**10**	**105**	**72**	**0,9162**
5	**49**	**11**	**121**	**72**	**0,8996**	**18**	**62**	**15**	**97**	**74**	**0,8710**
6	**47**	**14**	**114**	**81**	**0,9488**	**19**	**43**	**10**	**115**	**80**	**0,8287**
7	**30**	**13**	**112**	**72**	**0,9765**	**20**	**28**	**7**	**103**	**79**	**0,9330**
8	**46**	**14**	**119**	**81**	**0,9167**	**21**	**49**	**11**	**95**	**76**	**0,8805**
9	**35**	**9**	**121**	**72**	**0,8580**	**22**	**59**	**14**	**92**	**81**	**0,9696**
10	**53**	**8**	**118**	**79**	**0,8724**	**23**	**37**	**14**	**114**	**70**	**0,8332**
11	**28**	**10**	**102**	**81**	**0,9131**	**24**	**48**	**8**	**99**	**77**	**0,8147**
12	**31**	**12**	**92**	**78**	**0,8721**	**25**	**26**	**10**	**95**	**71**	**0,9642**
13	**54**	**15**	**119**	**80**	**0,9359**	**26**	**57**	**13**	**110**	**80**	**0,9398**

Оценка генеральной доли

Пусть из генеральной совокупности объёма N проведена выборка объёмом n, и по её результатам требуется оценить генеральную *долю* объектов, обладающих некоторым количественным или качественным признаком.

Пусть N –количество помидоров на базе, среди которых K первосортных.

Тогда отношение $\omega_Г=K/N$ является ***генеральной долей*** первосортных помидоров.

Однако исследовать все овощи затруднительно, поэтому организуется *представительная* выборка из n помидоров, среди которых первосортных окажется k штук.

Отношение $\omega_в=k/n$ называется ***выборочной долей***.

Задача состоит в том, чтобы по найденному значению $\omega_в$ оценить истинную долю $\omega_Г$.

Как оценить? С помощью доверительного интервала:

$(\omega_в - \Delta < \overline{x}_Г < \omega_в + \Delta)$, где Δ – ***предельная ошибка*** *доли*.

Задание 3

В целях изучения суточного пробега автомобилей автотранспортного предприятия проведено 10%-ное выборочное обследование 100 автомобилей методом случайного бесповторного отбора, в результате которого получены следующие данные:

Суточный пробег автомобиля, км	Число автомобилей
до 160	12
от 160 до 180	36
от 180 до 200	28
свыше 200	24
Итого	**100**

С вероятностью 0,954 требуется определить *долю* машин в генеральной совокупности с пробегом более 180 км.

Методика выполнения

Вычислим количество автомобилей с пробегом более 180 км по ***k***=28+24=52 выборке.

Таким образом: ω=k/n=52/100=0,52 – выборочная доля автомобилей с пробегом более 180 километров.

Генеральную долю $\boldsymbol{\omega_Г}$ таких автомобилей оценим с помощью доверительного интервала: $(\omega_в - \Delta < \overline{\omega_Г} < \omega_в + \Delta)$,

где ***Δ=t•μ*** – предельная ошибка доли (*коэффициента доверия* **t** на ***среднюю ошибку*** выборки ***μ***).

Для уровня доверительной вероятности *γ*=0,954 находим коэффициент доверия: Φ(t)=0,954/2=0,477 ⇒ ***t***=2.

Вычислим среднюю ошибку доли. Коль скоро, выборка 10%-ная, то объём генеральной совокупности равен ***N***= 100%/10%****n***= 1000 автомобилей и для бесповторной выборки:

$$\mu = \sqrt{\frac{\omega(1-\omega)}{n}\left(1-\frac{n}{N}\right)} = \sqrt{\frac{0{,}52*0{,}48}{100}\left(1-\frac{100}{1000}\right)} \approx 0{,}047.$$

Таким образом, предельная ошибка доли ***Δ=t•μ*** ≈ 2*0,047=0,095 и искомый доверительный интервал:

$$(\omega_в - \Delta < \omega_Г < \omega_в + \Delta)$$

$$0{,}52 - 0{,}095 < \omega_Г < 0{,}52 + 0{,}095$$

$$0{,}425 < \omega_Г < 0{,}615$$

Т.е. с вероятностью 95,4% данный интервал накрывает истинную генеральную долю $\boldsymbol{\omega_Г}$ автомобилей с пробегом более 180 км.

Ответ: $0{,}425 < \omega_Г < 0{,}615$.

Кстати, тут можно оценить и абсолютное количество таковых машин:

0,425***N***<***K***<0,615***N*** ⇒ 425<***K***<615.

Но результат, конечно, такой слабоватый. И помочь здесь может увеличение выборки.

Родственная формула уже выведена в предыдущем параграфе, и просто заменим дисперсию произведением ω(1-ω):

$$\boldsymbol{n = \frac{\omega(1-\omega)t^2 N}{\omega(1-\omega)t^2 + \Delta^2 N}}$$

– здесь по желаемой предельной ошибке можно вычислить необходимый объём выборки.

Варианты задания 3

№ вар.	*Суточ. пробег, км*					% выборки	*Пробег выше, км*	$P_{доверит.}$
	до 160	*от 160 до180*	*от180 до 200*	*от 200 до 220*	*свыше 220*			
	Число автом. в выборке, шт.							
1	12	19	16	19	12	12	192	0,8347
2	24	21	13	22	20	11	196	0,8595
3	21	13	23	18	18	12	192	0,8659
4	22	20	22	16	24	8	209	0,8933
5	18	16	14	19	20	11	208	0,8587
6	19	12	17	10	18	10	194	0,9751
7	22	16	16	24	20	12	214	0,9578
8	14	19	23	12	17	11	213	0,8316
9	10	22	17	22	12	12	181	0,9773
10	13	11	14	23	15	15	188	0,8004
11	20	21	19	18	20	11	206	0,9168
12	12	22	11	16	22	9	183	0,8254
13	17	15	11	20	15	10	200	0,8517
14	22	11	11	10	14	11	182	0,8989
15	23	14	14	19	16	10	192	0,8606
16	20	20	10	22	11	12	208	0,9199
17	20	19	14	21	14	12	190	0,9464
18	24	15	19	24	17	11	198	0,9317
19	20	19	21	12	22	13	181	0,9122
20	22	14	13	24	11	10	186	0,8030
21	22	18	14	21	20	10	208	0,9664
22	10	10	14	17	23	13	182	0,9509
23	16	20	16	22	13	14	193	0,8702
24	17	23	16	21	20	8	187	0,9646
25	23	23	11	22	14	8	203	0,8994

Задание 4

Методом механического отбора проведено 1%-ое обследование веса пирожных, изготовленных кондитерской фабрикой за сутки. Распределение веса пирожных по весу следующее:

Вес пирожных, г	96-98	98-100	100-102	102-104	Итого
Число пирожных	5	30	60	5	100

а) С вероятностью 0,9973 определить пределы, в которых будет находиться доля пирожных весом не менее 100 г, во всей суточной продукции?

б) Сколько процентов пирожных нужно проверить, чтобы улучшить оценку в 7 раз? (при той же доверительной вероятности)

Методика выполнения

а) Вычислим количество пирожных весом не менее 100 грамм

k=60+5=65.

Таким образом: ω=k/n=0,65 – выборочная доля таковых пирожных.

Соответствующую генеральную долю оценим с помощью доверительного интервала:

$(\omega_в - \Delta < \omega_Г < \omega_в + \Delta)$, где ***Δ=t•μ*** – предельная ошибка доли (*коэффициента доверия* ***t*** на ***средню ошибку*** выборки ***μ***).

Уровню доверительной вероятности γ=0,9973 соответствует коэффициент ***t***=3.

Вычислим среднюю ошибку доли. Поскольку выборка 1%-ная и бесповторная, то:

$$\mu = \sqrt{\frac{\omega(1-\omega)}{n}\left(1-\frac{n}{N}\right)} = \sqrt{\frac{0{,}65*0{,}35}{100}\left(1-\frac{1}{100}\right)} \approx 0{,}04746.$$

Таким образом, предельная ошибка доли ***Δ=t•μ***=3*0,04776 ≈ 0,14 и искомый доверительный интервал:

$$(\omega_в - \Delta < \omega_Г < \omega_в + \Delta)$$

$$0{,}65 - 0{,}14 < \omega_Г < 0{,}65 + 0{,}14$$

$$0{,}51 < \omega_Г < 0{,}79$$

– данный интервал практически достоверно накрывает долю пирожных весом не менее 100 грамм во всей суточной партии.

б) Улучшим точность оценки в 7 раз: ***Δ****=Δ/7=0,14/7=0,02 и вычислим объём выборки, которую следует организовать, чтобы обеспечить эту точность. Учитывая, что объём генеральной совокупности составляет ***N***=100%/1%****n***= 100*100=10000:

$$n^* = \frac{\omega(1-\omega)t^2 N}{\omega(1-\omega)t^2 + (\Delta^*)^2 N} = \frac{0,65*0,35*3^2*10000}{0,65*0,35*3^2 + 0,02^2*10000} \approx 3386$$

Таким образом, для того, чтобы с вероятностью 99,73% можно было утверждать, что выборочная доля $\boldsymbol{\omega^*}$ пирожных весом не менее 100 грамм будет отличаться от истинного значения $\boldsymbol{\omega_Г}$ менее чем на 0,02, следует организовать выборку объёмом $\boldsymbol{n}$=3386 пирожных, что составляет примерно треть генеральной совокупности.

Ответ: а) 0,51<$\boldsymbol{\omega_Г}$< 0,79 б) $\boldsymbol{n}$=3386.

Варианты задания 4

№ вар.	*Вес пирожных, г*					% выборки	*Вес не менее г*	$P_{доверит.}$
	до 96	*от 96 до98*	*от98 до 100*	*от 100 до 102*	*свыше 102*			
	Число автом. в выборке, шт.							
1	11	23	49	41	6	4	100	0,8282
2	5	19	55	54	14	2	100	0,8387
3	12	28	42	27	8	5	99	0,9002
4	12	23	47	48	6	4	100	0,9425
5	15	16	46	51	13	2	99	0,9776
6	14	30	33	41	15	1	98	0,8632
7	9	30	44	29	6	5	101	0,9264
8	14	22	50	46	5	3	100	0,8067
9	13	19	28	49	6	5	100	0,8226
10	13	23	26	47	11	3	101	0,8957
11	5	30	26	41	12	5	98	0,9339
12	15	26	38	45	12	4	101	0,8813
13	8	35	26	31	12	4	101	0,8818
14	10	33	49	44	13	2	101	0,8416
15	11	28	55	42	5	2	101	0,8427
16	12	25	27	32	5	2	98	0,9489
17	11	33	25	45	13	3	101	0,9742
18	12	23	32	36	12	5	101	0,9096
19	9	33	54	25	14	3	98	0,9531
20	8	18	50	55	12	1	100	0,9546
21	12	24	36	48	5	3	99	0,9422
22	6	24	41	28	15	1	101	0,9344
23	6	19	46	35	12	5	98	0,9635
24	6	22	34	43	15	1	98	0,8320
25	10	21	51	38	6	5	101	0,9692