

Chap 3r Sampling Distribution

(Parameter)

Population Population is the aggregate or total of statistical data forming a subject of investigation.

e.g. population of workers in a factory, population of nationalized bank in India etc.

→ The no. of observations in a population is defined to be size of population and is denoted by ' N '.

→ If the no. of units in the population is finite then the population is called finite population.
e.g. The population of students in a college.

→ If the size of population is infinite then it is called infinite population.

e.g. The population of stars in the sky.

Sample Most of the times study of entire population

may not be possible in such cases we can select some units from the population i.e. a part (or) subset of population which is examined with a view of to determine the population characteristic is called a sample.

→ The no. of obs in a sample is called sample size and is denoted by ' n '.

→ If the sample size ' n ' is greater than or equal to 30 then the sample is called large sample.

→ If the sample size 'n' is less than 30 then the sample is called small sample.

Sampling Techniques or Sampling methods: The technique of drawing samples from the population is called Sampling Methods. Some important methods of sampling are discussed below.

(1) Simple Random Sampling Method (SRS): It is the process of

drawing a sample from the population in such a way that each and every unit of population has an equal chance of being included in the sample. The sample obtained by this process is called random sample.

(2) If each element of a population may be selected more than once then it is called simple random sampling with replacement whereas if the element cannot be selected more than once then it is called simple random sampling without replacement.

→ The no. of samples with replacement is N^n .

→ The no. of samples without replacement is ${}^N C_n$,

where $N =$ population size
 $n =$ sample size

(2) Stratified Sampling: This method is useful when the population is heterogeneous. In this type of sampling the population is first subdivided into several groups. These groups are called strata. Then a sample is selected from each stratum at random. The process of obtaining and examining from stratified

- 3 Sample with a view to estimating the characteristics of the population is known as stratified random sample.
- 4 ③ Systematic Sampling Method In this method all the units of the population are arranged in some order. Then from the first k items, one unit is selected at random. By continuing this process we get systematic sample.
- 5 Parameter: It is a statistical measure based on all the observations of population.

Population mean μ , Population standard deviation σ , Population proportion p etc are usually referred as a parameter.

Statistic: It is a statistical measure based only on all the units selected in a sample.

Sample mean \bar{x} , sample S.D's, sample proportion 'p' are usually referred to as statistics.

Sampling distribution: sampling theory is the study of relationship b/w a population and samples drawn from the population.

Sampling distribution of a statistic is the frequency distribution which is informed with various values of a statistic computed from diff. samples of the same size drawn from the same population. These various values can be arranged into a frequency distribution table which is known as the Sampling distribution of the statistic.

Let us consider a finite population of size N and drawn all possible random samples of each of same size n , then we get k samples. Now the sampling distribution of mean and standard deviation can be tabulated as follows

Samples	Sample units	Sample mean	Sample S.D.
1	x_1, x_2, \dots, x_{1n}	\bar{x}_1	$s_1 = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x}_1)^2}$
2	$x_2, x_{22}, \dots, x_{2n}$	\bar{x}_2	$s_2 = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x}_2)^2}$
i	\vdots	\vdots	\vdots
k	$x_k, x_{k2}, \dots, x_{kn}$	\bar{x}_k	$s_k = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x}_k)^2}$

Standard error of a statistic The standard error of a statistic is the standard deviation of the sampling distribution of statistic

formulas for standard error

$$① \text{ Standard error of sample mean } S.E(\bar{x}) = \sqrt{\frac{\sigma^2}{n}}$$

$$② S.E \text{ of the diff of 2 sample mean } = S.E(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$③ S.E \text{ of sample S.D. } S = S.E(S) = \sqrt{\frac{\sigma^2}{2n}}$$

$$④ S.E \text{ of sample proportion } p = S.E(p) = \sqrt{\frac{pq}{n}}$$

$$⑤ S.E \text{ of diff of two proportions } = S.E(p_1 - p_2) = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

Note Suppose the samples are drawn from an infinite population i.e. Sampling is done with replacement then a sampling distribution of

3 Mean is equal to population mean i.e. $\bar{M} = \mu$

$$\text{Variance}(\bar{x}) = \frac{\sigma^2}{n}$$

4 If the samples are drawn from a finite population i.e. sampling is done without replacement then the mean of sampling distribution is = population mean

i.e. $\bar{M} = \mu$

$$V(\bar{x}) = \left(\frac{N-n}{N-1} \right) \cdot \frac{\sigma^2}{n}$$

here $\frac{N-n}{N-1}$ is known as Finite Population correction factor

Prob

1. Find the value of finite population correction factor for $n=10$ & $N=100$

$$N=100$$

$$\left(\frac{N-n}{N-1} \right) = \left(\frac{100-10}{100-1} \right) = \frac{90}{99} = 0.909$$

2. A population consists of 5 numbers 2, 3, 6, 8, 11 consider all possible samples of size 2 which can be drawn with replacement from this population. Find (1) The population mean (2) The S.D. of population (3) The mean of sampling distribution of means (4) The S.D. of sampling distribution of means

Sol

$$N=5, n=2$$

(1) population mean = $\mu = \frac{2+3+6+8+11}{5} = 6.8$

individual probabilities of each sample sum up to 1

$$\textcircled{2} \quad r = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} = \sqrt{\frac{(2-6)^2 + (3-6)^2 + (6-6)^2 + (8-6)^2 + (11-6)^2}{5}} = \sqrt{18.2863}$$

The total no. of samples with replacement is $N^n = 5^2 = 25$ samples

(2,2) (2,3) (2,6) (2,8) (2,11)

$(3, 2)$ $(3, 3)$ $(3, 6)$ $(3, 8)$ $(3, 11)$

$(6, 2)$ $(6, 3)$ $(6, 6)^T$. $(6, 8) \oplus (6, 10)$, 3. 16. 2018 10:30 AM

(8,3) (8,6) (8,8) (8,11) 3 812 6 1013 8 1114 10 1215 12 1316 14 1517 16 1718 18 1919 20 2122 22 2324 24 2526 26 2728 28 2929 30 3132 32 3334 34 3536 36 3738 38 3939 40 4140 40 4243 42 4344 44 4545 45 4646 46 4747 47 4848 48 4949 49 5050 50 5151 51 5252 52 5353 53 5454 54 5555 55 5656 56 5757 57 5858 58 5959 59 6060 60 6161 61 6262 62 6363 63 6464 64 6565 65 6666 66 6767 67 6868 68 6969 69 7070 70 7171 71 7272 72 7373 73 7474 74 7575 75 7676 76 7777 77 7878 78 7979 79 8080 80 8181 81 8282 82 8383 83 8484 84 8585 85 8686 86 8787 87 8888 88 8989 89 9090 90 9191 91 9292 92 9393 93 9494 94 9595 95 9696 96 9797 97 9898 98 9999 99 100100 100 101101 101 102102 102 103103 103 104104 104 105105 105 106106 106 107107 107 108108 108 109109 109 110110 110 111111 111 112112 112 113113 113 114114 114 115115 115 116116 116 117117 117 118118 118 119119 119 120120 120 121121 121 122122 122 123123 123 124124 124 125125 125 126126 126 127127 127 128128 128 129129 129 130130 130 131131 131 132132 132 133133 133 134134 134 135135 135 136136 136 137137 137 138138 138 139139 139 140140 140 141141 141 142142 142 143143 143 144144 144 145145 145 146146 146 147147 147 148148 148 149149 149 150150 150 151151 151 152152 152 153153 153 154154 154 155155 155 156156 156 157157 157 158158 158 159159 159 160160 160 161161 161 162162 162 163163 163 164164 164 165165 165 166166 166 167167 167 168168 168 169169 169 170170 170 171171 171 172172 172 173173 173 174174 174 175175 175 176176 176 177177 177 178178 178 179179 179 180180 180 181181 181 182182 182 183183 183 184184 184 185185 185 186186 186 187187 187 188188 188 189189 189 190190 190 191191 191 192192 192 193193 193 194194 194 195195 195 196196 196 197197 197 198198 198 199199 199 200200 200 201201 201 202202 202 203203 203 204204 204 205205 205 206206 206 207207 207 208208 208 209209 209 210210 210 211211 211 212212 212 213213 213 214214 214 215215 215 216216 216 217217 217 218218 218 219219 219 220220 220 221221 221 222222 222 223223 223 224224 224 225225 225 226226 226 227227 227 228228 228 229229 229 230230 230 231231 231 232232 232 233233 233 234234 234 235235 235 236236 236 237237 237 238238 238 239239 239 240240 240 241241 241 242242 242 243243 243 244244 244 245245 245 246246 246 247247 247 248248 248 249249 249 250250 250 251251 251 252252 252 253253 253 254254 254 255255 255 256256 256 257257 257 258258 258 259259 259 260260 260 261261 261 262262 262 263263 263 264264 264 265265 265 266266 266 267267 267 268268 268 269269 269 270270 270 271271 271 272272 272 273273 273 274274 274 275275 275 276276 276 277277 277 278278 278 279279 279 280280 280 281281 281 282282 282 283283 283 284284 284 285285 285 286286 286 287287 287 288288 288 289289 289 290290 290 291291 291 292292 292 293293 293 294294 294 295295 295 296296 296 297297 297 298298 298 299299 299 300300 300 301301 301 302302 302 303303 303 304304 304 305305 305 306306 306 307307 307 308308 308 309309 309 310310 310 311311 311 312312 312 313313 313 314314 314 315315 315 316316 316 317317 317 318318 318 319319 319 320320 320 321321 321 322322 322 323323 323 324324 324 325325 325 326326 326 327327 327 328328 328 329329 329 330330 330 331331 331 332332 332 333333 333 334334 334 335335 335 336336 336 337337 337 338338 338 339339 339 340340 340 341341 341 342342 342 343343 343 344344 344 345345 345 346346 346 347347 347 348348 348 349349 349 350350 350 351351 351 352352 352 353353 353 354354 354 355355 355 356356 356 357357 357 358358 358 359359 359 360360 360 361361 361 362362 362 363363 363 364364 364 365365 365 366366 366 367367 367 368368 368 369369 369 370370 370 371371 371 372372 372 373373 373 374374 374 375375 375 376376 376 377377 377 378378 378 379379 379 380380 380 381381 381 382382 382 383383 383 384384 384 385385 385 386386 386 387387 387 388388 388 389389 389 390390 390 391391 391 392392 392 393393 393 394394 394 395395 395 396396 396 397397 397 398398 398 399399 399 400400 400 401401 401 402402 402 403403 403 404404 404 405405 405 406406 406 407407 407 408408 408 409409 409 410410 410 411411 411 412412 412 413413 413 414414 414 415415 415 416416 416 417417 417 418418 418 419419 419 420420 420 421421 421 422422 422 423423 423 424424 424 425425 425 426426 426 427427 427 428428 428 429429 429 430430 430 431431 431 432432 432 433433 433 434434 434 435435 435 436436 436 437437 437 438438 438 439439 439 440440 440 441441 441 442442 442 443443 443 444444 444 445445 445 446446 446 447447 447 448448 448 449449 449 450450 450 451451 451 452452 452 453453 453 454454 454 455455 455 456456 456 457457 457 458458 458 459459 459 460460 460 461461 461 462462 462 463463 463 464464 464 465465 465 466466 466 467467 467 468468 468 469469 469 470470 470 471471 471 472472 472 473473 473 474474 474 475475 475 476476 476 477477 477 478478 478 479479 479 480480 480 481481 481 482482 482 483483 483 484484 484 485485 485 486486 486 487487 487 488488 488 489489 489 490490 490 491491 491 492492 492 493493 493 494494 494 495495 495 496496 496 497497 497 498498 498 499499 499 500500 500 501501 501 502502 502 503503 503 504504 504 505505 505 506506 506 507507 507 508508 508 509509 509 510510 510 511511 511 512512 512 513513 513 514514 514 515515 515 516516 516 517517 517 518518 518 519519 519 520520 520 521521 521 522522 522 523523 523 524524 524 525525 525 526526 526 527527 527 528528 528 529529 529 530530 530 531531 531 532532 532 533533 533 534534 534 535535 535 536536 536 537537 537 538538 538 539539 539 540540 540 541541 541 542542 542 543543 543 544544 544 545545 545 546546 546 547547 547 548548 548 549549 549 550550 550 551551 551 552552 552 553553 553 554554 554 555555 555 556556 556 557557 557 558558 558 559559 559 560560 560 561561 561 562562 562 563563 563 564564 564 565565 565 566566 566 567567 567 568568 568 569569 569 570570 570 571571 571 572572 572 573573 573 574574 574 575575 575 576576 576 577577 577 578578 578 579579 579 580580 580 581581 581 582582 582 583583 583 584584 584 585585 585 586586 586 587587 587 588588 588 589589 589 590590 590 591591 591 592592 592 593593 593 594594 594 595595 595 596596 596 597597 597 598598 598 599599 599 600600 600 601601 601 602602 602 603603 603 604604 604 605605 605 606606 606 607607 607 608608 608 609609 609 610610 610 611611 611 612612 612 613613 613 614614 614 615615 615 616616 616 617617 617 618618 618 619619 619 620620 620 621621 621 622622 622 623623 623 624624 624 625625 625 626626 626 627627 627 628628 628 629629 629 630630 630 631631 631 632632 632 633633 633 634634 634 635635 635 636636 636 637637 637 638638 638 639639 639 640640 640 641641 641 642642 642 643643 643 644644 644 645645 645 646646 646 647647 647 648648 648 649649 649 650650 650 651651 651 652652 652 653653 653 654654 654 655655 655 656656 656 657657 657 658658 658 659659 659 660660 660 661661 661 662662 662 663663 663 664664 664 665665 665 666666 666 667667 667 668668 668 669669 669 670670 670 671671 671 672672 672 673673 673 674674 674 675675 675 676676 676 677677 677 678678 678 679679 679 680680 680 681681 681 682682 682 683683 683 684684 684 685685 685 686686 686 687687 687 688688 688 689689 689 690690 690 691691 691 692692 692 693693 693 694694 694 695695 695 696696 696 697697 697 698698 698 699699 699 700700 700 701701 701 702702 702 703703 703 704704 704 705705 705 706706 706 707707 707 708708 708 709709 709 710710 710 711711 711 712712 712 713713 713 714714 714 715715 715 716716 716 717717 717 718718 718 719719 719 720720 720 721721 721 722722 722 723723 723 724724 724 725725 725 726726 726 727727 727 728728 728 729729 729 730730 730 731731 731 732732 732 733733 733 734734 734 735735 735 736736 736 737737 737 738738 738 739739 739 740740 740 741741 741 742742 742 743743 743 744744 744 745745 745 746746 746 747747 747 748748 748 749749 749 750750 750 751751 751 752752 752 753753 753 754754 754 755755 755 756756 756 757757 757 758758 758 759759 759 760760 760 761761 761 762762 762 763763 763 764764 764 765765 765 766766 766 767767 767 768768 768 769769 769 770770 770 771771 771 772772 772 773773 773 774774 774 775775 775 776776 776 777777 777 778778 778 779779 779 780780 780 781781 781 782782 782 783783 783 784784 784 785785 785 786786 786 787787 787 788788 788 789789 789 790790 790 791791 791 792792 792 793793 793 794794 794 795795 795 796796 796 797797 797 798798 798 799799 799 800800 800 801801 801 802802 802 803803 803 804804 804 805805 805 806806 806 807807 807 808808 808 809809 809 810810 810 811811 811 812812 812 813813 813 814814 814 815815 815 816816 816 817817 817 818818 818 819819 819 820820 820 821821 821 822822 822 823823 823 824824 824 825825 825 826826 826 827827 827 828828 828 829829 829 830830 830 831831 831 832832 832 833833 833 834834 834 835835 835 836836 836 837837 837 838838 838 839839 839 840840 840 841841 841 842842 842 843843 843 844844 844 845845 845 846846 846 847847 847 848848 848 849849 849 850850 850 851851 851 852852 852 853853 853 854854 854 855855 855 856856 856 857857 857 858858 858 859859 859 860860 860 861861 861 862862 862 863863 863 864864 864 865865 865 866866 866 867867 867 868868 868 869869 869 870870 870 871871 871 872872 872 873873 873 874874 874 875875 875 876876 876 877877 877 878878 878 879879 879 880880 880 881881 881 882882 882 883883 883 884884 884 885885 885 886886 886 887887 887 888888 888 889889 889 890890 890 891891 891 892892 892 893893 893 894894 894 895895 895 896896 896 897897 897 898898 898 899899 899 900900 900 901901 901 902902 902 903903 903 904904 904 905905 905 906906 906 907907 907 908908 908 909909 909 910910 910 911911 911 912912 912 913913 913 914914 914 915915 915 916916 916 917917 917 918918 918 919919 919 920920 920 921921 921 922922 922 923923 923 924924 924 925925 925 926926 926 927927 927 928928 928 929929 929 930930 930 931931 931 932932 932 933933 933 934934 934 935935 935 936936 936 937937 937 938938 938 939939 939 940940 940 941941 941 942942 942 943943 943 944944 944 945945 945 946946 946 947947 947 948948 948 949949 949 950950 950 951951 951 952952 952 953953 953 954954 954 955955 955 956956 956 957957 957 958958 958 959959 959 960960 960 961961 961 962962 962 963963 963 964964 964 965965 965 966966 966 967967 967 968968 968 969969 969 970970 970 971971 971 972972 972 973973 973 974974 974 975975 975 976976 976 977977 977 978978 978 979979 979 980980 980 981981 981 982982 982 983983 983 984984 984 985985 985 986986 986 987987 987 988988 988 989989 989 990990 990 991991 991 992992 992 993993 993 994994 994 995995 995 996996 996 997997 997 998998 998 999999 999

(11,3) (11,6) (11,8) (11,11) \rightarrow 10'chen
11,2 11,3 11,4 11,5 11,6 11,7 11,8 11,9 11,10 11,11

sample mean(\bar{x}) = $\frac{1}{n} \sum x_i$; $\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n)$

2.5, 3, 4.5, 5.5, 7

4, 4.5, 6, 7, 8.5 = 150
3.50 2.50

$$= 11 = 5, -5, 7, 8, 9, \text{ so } 11 \text{ is a right neighbor of } 5$$

6.5, 7, 8.5, 9.5, 11

$$(x-\mu)^2 = 16, 12.25, 10, 4, 1, 0$$

62.25 9 2.25, 0.25, 1

12.25, 88FP-1 6.25

4, 2.25, 0,

loss of life began January 1, 1945, 12:25

1, 0.25, 12:25 325

12.25 6.25 , 12.25

0.25, , ,

mean of sampling distribution of means $\bar{M}_{\bar{x}} = \frac{\sum x}{n} = \frac{100}{25} = 4$

③ The mean of

Sampling distrib

نَفَرَتْ مُهَاجِرَةً فَعَلَّمَهُ

W.S.V.W

④ The S.D. of sampling distribution of means $\sigma_{\bar{x}} = \sqrt{\frac{\sum (x_i - \mu)^2}{n}}$

$$= \sqrt{\frac{135}{25}} = 2.3238$$

$\therefore \sigma_{\bar{x}} = \sqrt{\frac{\sum (x_i - \mu)^2}{n}}$

Prob

3. A population consists of 5, 10, 14, 18, 13, 24. Consider all possible samples of size 2 which can be drawn without replacement from this population. Then find
- ① Population mean
 - ② S.D. of population
 - ③ The mean of sampling distribution of means
 - ④ S.D. of sampling distribution of means

Sol

$$N = 6, n = 2$$

$$\textcircled{1} \text{ Population mean } \mu = \frac{5+10+14+18+13+24}{6} = 14$$

$$\textcircled{2} \sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}} = \sqrt{\frac{(5-14)^2 + (10-14)^2 + (14-14)^2 + (18-14)^2 + (13-14)^2 + (24-14)^2}{6}} = 8.9722$$

- ⑤ In case of without replacement method the total possible samples = $N C_2 = 15$

⑥ Samples

$$\begin{array}{l} (5, 10) : (5, 14) : (5, 18) : (5, 13) : (5, 24) \\ (10, 14) : (10, 18) : (10, 13) : \cancel{(10, 24)} : (10, 24) \\ (14, 18) : (14, 13) : (14, 24) \end{array}$$

(18, 13) (18, 24)

(13, 24)

Sample means

7.5, 9.5, 11.5, 9.5, 14.5, members of ①

12, 14, 11.5, 17

16, 13.5, 19.5, 13.5, 21.5, $\bar{x} = 21.0$

15.5, 21, 2

18.5

members of ②

$$(\bar{x} - \mu)^2 = 4.25$$

$$20.25, 6.25, 25, 0.25, 3.25, 11$$

$$4, 0, 6.25, 9, 12, 22, 3.25, 11$$

$$4, 0.25, 25, 0.25, 214$$

$$(15, 21, 8), (15, 7, 3), (18, 9), (15, 3, 2), (21, 3, 9), (8, 3, 2)$$

$$2.25, 9$$

$$(15, 21, 8), (15, 7, 3), (18, 9)$$

$$20.25$$

$$\bar{x} = \frac{\sum \bar{x}}{15} = \frac{210}{15}$$

③ The mean of sampling distribution of mean $\bar{M}_x = \frac{\sum \bar{x}}{15} = \frac{210}{15} = 14$

$$\bar{M}_x = 14$$

④ The S.D of sampling distribution of mean $\sigma_x = \sqrt{\frac{\sum (\bar{x} - \bar{M}_x)^2}{15}}$

$$= \sqrt{\frac{214}{15}} = 3.7721$$

$$\therefore \sigma \neq \sigma_x$$

4. If the population is 3, 6, 9, 15, 27 list all possible samples of size 3 that can be taken without replacement from this population
① Mean of population ② S.D of population
③ The mean of a sampling distribution of means ④ S.D of sampling distribution of means

SOL

④ $N = 5, n = 3$

$$\textcircled{1} \text{ Population mean } = \mu_p = \frac{3+6+9+15+27}{5} = 12.1$$

+ mean sigma

$$\textcircled{2} \sigma = \sqrt{\frac{\sum (x-\mu)^2}{N}} = \sqrt{\frac{(3-12)^2 + (6-12)^2 + (9-12)^2 + (15-12)^2 + (27-12)^2}{5}} = 8.48583$$

In case of without replacement the total possible samples

$$= Ncn = 5C_3 = \frac{5!}{(5-3)!3!} = 10$$

- samples
- (3, 6, 9) (3, 6, 15) (3, 6, 27) (3, 9, 15) (3, 9, 27) (3, 15, 27)
 - (6, 9, 15) (6, 9, 27) (6, 15, 27)
 - (9, 15, 27)

$$\bar{x} = \frac{3+6+9+15+27}{5} = 12$$

$$(x-\bar{x})^2 = 36 \quad 16 \quad 0 \quad 9 \quad 1 \quad 9$$

13

15, 18, 21, 24, 27

$$(x-\bar{x})^2 = 36 \quad 16 \quad 0 \quad 9 \quad 1 \quad 9$$

4 4 16

$$= 120$$

Mean of sum = $\frac{120}{10} = 12$

$$\textcircled{3} : \mu_{\bar{x}} = \frac{\sum x}{n} = \frac{120}{10} = 12$$

$$\mu_x = 12$$

$$(k) \sigma_{\bar{x}} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{10}} = \sqrt{\frac{120}{10}} = 3.604$$

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

Q. what is the effect on standard error on mean of the sample size is taken from an infinite population if sample size is increased from 400 - 900

$$n_1 = 400; n_2 = 900$$

~~work~~ $s.e(\bar{x}) = \sigma / \sqrt{n}$

$$[s.e(\bar{x})]_{n_1} = \sigma / \sqrt{n_1} = \frac{\sigma}{\sqrt{400}} = \frac{\sigma}{20}$$

$$[s.e(\bar{x})]_{n_2} = \frac{\sigma}{\sqrt{n_2}} = \frac{\sigma}{\sqrt{900}} = \frac{\sigma}{30}$$

$$[s.e(\bar{x})]_{n_1} = \frac{\sigma}{20} = \frac{\sigma}{30} \times \frac{3}{2}$$

$$\text{standard deviation of } \bar{x} = \frac{3}{2} (\text{s.e}(\bar{x}))_{n_2}$$

$$[s.e(\bar{x})]_{n_1} = \frac{3}{2} \times [s.e(\bar{x})]_{n_2}$$

Q. When the sample is taken from an infinite population what happened to standard error of mean if the sample size is decreased from 800 to 200

$$n_1 = 800, n_2 = 200$$

$$s.e(\bar{x}) = \frac{\sigma}{\sqrt{n}}$$

$$S.E(\bar{x})]_{n_1} = \frac{\sigma}{\sqrt{n}} = \frac{\sigma}{\sqrt{800}} = \frac{\sigma}{20\sqrt{2}} = \frac{0.5}{20\sqrt{2}} = \frac{0.5}{28.28} = 0.0178$$

$$S.E(\bar{x})]_{n_2} = \frac{\sigma}{\sqrt{n}} = \frac{\sigma}{\sqrt{1000}} = 0.0178$$

$$= 2 \cdot \frac{\sigma}{20\sqrt{2}} = 2, S.E(\bar{x})]_{n_1}$$

Since σ is to remain the same in both cases \Rightarrow

$$S.E(\bar{x})]_{n_2} = 2 S.E(\bar{x})]_{n_1}$$

Estimation

Estimator An estimate is a statement made to find an unknown population parameter

Estimator The procedure to determine an unknown population parameter is called estimator

Types of estimation - Basically there are two kinds of estimates to determine the statistic of population parameters namely point estimation & interval estimation

Point estimation If an estimate of population parameter is given by a single value then the estimate is called a point estimation

Properties of good estimator A good estimator is one which is closed to the true value of parameter. The properties of a good estimator are unbiasedness, consistency and efficiency.

Unbiased estimator A statistic is said to be an unbiased estimator of the corresponding parameter if the mean of sampling distribution of the statistic is equal to the corresponding population parameter otherwise the statistic is called a biased estimator of the corresponding parameter if t is a statistic and θ be the corresponding parameter and if $E(t) = \theta \Rightarrow E[\text{statistic}] = \text{Parameter}$ the 't' is an unbiased estimator of θ .

Consistency estimator An estimator t of a parameter θ is said to be consistent if it satisfies the following 2 conditions.

$$E(t) = \theta \text{ & } V(t) \rightarrow 0 \text{ as } n \rightarrow \infty$$

Efficient estimator A statistic t_1 is said to be more efficient unbiased estimator of the parameter θ than the statistic t_2 if it satisfies the following 2 conditions.

① t_1 & t_2 are both unbiased estimators of θ

② $V(t_1) < V(t_2)$

Max error of an estimate for large samples

Max error of an estimate of the population mean in case of large sample is given by

$$\text{Max error}(ME) = t(\alpha/2) \left[\frac{\sigma}{\sqrt{n}} \right]$$

In case of small samples the max error of an estimate is given by

$$\text{Max error}(ME) = t(\alpha/2) (s/\sqrt{n})$$

Here $Z(\alpha/2)$ is the table values of Z from the normal tables.

n size of samples

σ population s.d.

$t(\alpha/2)$ is the table values of t distribution critical values

s is s.d. of samples

* Max error of proportion is $M.E = Z(\alpha/2) \cdot \sqrt{\frac{pq}{n}}$

where $Q = 1 - P$

Here P is proportion of success of an event

* In case of small samples the max error of population mean is given by $M.E = t(\alpha/2) \cdot \sqrt{\frac{S^2}{n}}$

* Interval estimation - An estimate of a population parameter is given by two diff values b/w which the parameter (may be, considered to lies then the estimate is called an interval estimate of the parameter.)

For large samples the interval estimation of population mean is given by

$$\mu = (\bar{x} - Z(\alpha/2) \cdot \sqrt{\frac{s^2}{n}}) \text{ to } (\bar{x} + Z(\alpha/2) \cdot \sqrt{\frac{s^2}{n}})$$

The interval estimation of difference of two population mean is given by

$$[(\bar{x}_1 - \bar{x}_2) - Z(\alpha/2) \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}] \text{ to } [(\bar{x}_1 - \bar{x}_2) + Z(\alpha/2) \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}]$$

The interval estimation of population proportion (p) is given by

$$P = \left(P - 2(\alpha/2) \sqrt{\frac{PQ}{n}} ; P + 2\alpha/2 \sqrt{\frac{PQ}{n}} \right)$$

The interval estimation of diff of two population proportions $P_1 - P_2$

$$P_1 - P_2 = \left[(P_1 - P_2) \pm 2 \times 1/2 \sqrt{PQ \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \right]$$

where $P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$, $Q = 1 - P$

where \bar{x} = sample mean, $(x_1 + x_2 + \dots + x_n) / n$

σ = population standard deviation

$Z_{\alpha/2}$ = The table value of 'z' at $\alpha/2$ level from the normal tables

P = Sample proportion

p = population proportion

n = large sample size n_1, n_2 are also large sample sizes

In case of small samples [~~not~~] the interval estimation of population

Mean is given by

$$u = \left(\bar{x} - t e(\alpha_{12}) s/\sqrt{n}; \quad \bar{x} + t e(\alpha_{12}) s/\sqrt{n} \right) \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

The interval estimation of difference of two population means

$$\bar{x}_1 - \bar{x}_2 = \left((\bar{x}_1 - \bar{x}_2) - t(\alpha/2) \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right); \quad (\bar{x}_1 - \bar{x}_2) + t(\alpha/2) \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

where \bar{x} is the sample mean ~~$\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n-1}$~~

s - is the sample s.d

$t_{\alpha/2}$ = The table value of t at $\alpha/2$ level from the t table
 values where the population S.D can be estimated through the samples by using following formula

$$\boxed{s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2} \quad (\text{or}) \quad \frac{\sum (x_{ij} - \bar{x}_1)^2 + \sum (x_{2j} - \bar{x}_2)^2}{(n_1 + n_2) - 2}}$$

Some Results on Estimation:

1. S.T the sample mean \bar{x} is an unbiased estimator of the population mean μ

Proof Let x_1, x_2, \dots, x_n be the samples of size n drawn from the population with mean μ and variance σ^2

i.e expectation of x_i is $\mu \Rightarrow E(x_1) = \mu; E(x_2) = \mu; \dots; E(x_n) = \mu$

i.e $E(\bar{x}) = \text{Avg} = \text{mean}$

$$V(x_1) = \sigma^2; V(x_2) = \sigma^2, \dots; V(x_n) = \sigma^2$$

$$E(\text{statistic}) = E(\text{sample mean})$$

$$= E\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right)$$

$$= \frac{1}{n} [E(x_1) + E(x_2) + \dots + E(x_n)]$$

$$= \frac{1}{n} [\mu + \mu + \dots + \mu \text{ (n times)}]$$

$$= \frac{1}{n} \cdot n \cdot \mu = \mu = \text{population mean}$$

\therefore sample mean \bar{x} is an unbiased estimator of the population mean μ

2. S.T. for a random sample of size n , taken from an infinite population.

$s^2 = \frac{1}{n} \sum (x_i - \bar{x})^2$ is not an unbiased estimator of a population variance σ^2 but $s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$ is an unbiased estimator of Population Variance σ^2

Soln Let x_1, x_2, \dots, x_n be the samples of size n drawn from the population with mean μ variance σ^2

$$\text{i.e. } E(x_1) = \mu; E(x_2) = \mu; \dots E(x_n) = \mu$$

$$\text{i.e. } E(x) = \text{Avg} = \text{mean}$$

$$V(x_1) = \sigma^2; V(x_2) = \sigma^2; \dots V(x_n) = \sigma^2$$

$$\begin{aligned} \text{Consider } s^2 &= \frac{1}{n} \sum (x_i - \bar{x})^2 \\ &= \frac{1}{n} \left[\sum (x_i^2 + \bar{x}^2 - 2x_i \bar{x}) \right] \\ &= \frac{1}{n} \left[\sum x_i^2 + \sum \bar{x}^2 - 2\bar{x} \sum x_i \right] \\ &= \frac{1}{n} \left[\sum x_i^2 + n\bar{x}^2 - 2\bar{x} \cdot n\bar{x} \right] \quad \therefore \bar{x} = \frac{\sum x_i}{n} \\ &= \frac{1}{n} \left[\sum x_i^2 + n\bar{x}^2 - 2n\bar{x}^2 \right] \\ &= \frac{1}{n} \left[\sum x_i^2 - n\bar{x}^2 \right] \end{aligned}$$

$$= \frac{1}{n} \sum x_i^2 - \bar{x}^2$$

$$E(s^2) = E \left(\frac{1}{n} \sum x_i^2 \right) - E(\bar{x}^2)$$

$$= \frac{1}{n} \cdot E(\sum x_i^2) - E(\bar{x}^2)$$

Now w.k.t

$$V(x) = \frac{1}{n} \sum x_i^2 - \left(\frac{\sum x_i}{n} \right)^2$$

$$3. \quad \sigma^2 = [E(\bar{x})]^2$$

$$\therefore \sigma^2 = E(x_i^2) - \mu^2$$

$$4. \quad E(x_i^2) = \sigma^2 + \mu^2 \quad \text{--- (1)}$$

$$5. \quad \sigma^2 = E(x^2) - [E(\bar{x})]^2$$

$$\therefore \frac{\sigma^2}{n} = E(\bar{x}^2) - [E(\bar{x})]^2 = \mu^2$$

$$E(\bar{x}^2) = \frac{\sigma^2}{n} + \mu^2 \quad \text{--- (2)}$$

now substitute (1) & (2) in,

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

$$= \frac{1}{n} n \cdot E(x_i^2) - E(\bar{x}^2)$$

$$= E(x_i^2) - E(\bar{x}^2)$$

$$= \sigma^2 + \mu^2 - \frac{\sigma^2}{n} - \mu^2$$

$$= \sigma^2 \left(1 - \frac{1}{n}\right)$$

$$= \sigma^2 \left(\frac{n-1}{n}\right)$$

$\neq \sigma^2$

$\therefore s^2$ is not an unbiased estimator of σ^2

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

$$= \frac{n}{n-1} \times \frac{1}{n} \sum (x_i - \bar{x})^2$$

$$= \frac{n}{n-1} \cdot s^2$$

$$E(s^2) = \frac{n}{n-1} \cdot E(s^2)$$

$$= \frac{n}{n-1} \cdot \frac{s^2}{\sigma^2} = n - 1$$

s^2 is an unbiased estimator of σ^2

Note: In case of large samples the table of Z i.e. $Z_{\alpha/2}$ at 95% level is 1.96 & at 99% level is 2.58 (alpha level is 1.645)

Ques. A random sample of size 100 has a S.D. of 5 what can you say about the max error with 90 95% confidence & also construct the confidence limits with the sample mean \bar{x} is 23

Sol: Given that $n=100$ which is greater than or equal to 30 the prob is related to large sample prob.

$$S.D. = \sigma = 5$$

$$\text{at } 95\% \text{ confidence } Z_{\alpha/2} = 1.96$$

$$Z_{\alpha/2} = 1.96$$

$$\alpha = 95\%$$

$$\beta = 0.95$$

$$\frac{\alpha}{2} = \frac{0.95}{2} = 0.475$$

for large sample

$$M.E. \text{ of Actual Mean } \bar{x} = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$= 1.96 \times \frac{5}{\sqrt{100}} = 0.98$$

$$\text{Confidence in } 95\% \text{ for } \bar{x} = [\bar{x} - Z_{\alpha/2} (\sigma/\sqrt{n}), \bar{x} + Z_{\alpha/2} (\sigma/\sqrt{n})]$$

$$= 23 - 0.98 = 22.02$$

$$= 23 + 0.98 = 23.98$$

Q. Total of the max. error one can accept to above with the probability 0.90 when using the mean of a random sample of size 64 to estimate the mean of population with variance 2.86 and also construct confidence limits for it when the sample mean is 32.

$$N = 64 \Rightarrow n = 8$$

Q. If it's large sample then what is the result?

Given $\sigma^2 = 2.86$, $n = 64$, $\bar{x} = 32$, $1 - \alpha = 0.90$

$$E(\bar{x}) = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{2.86}{64}} = 0.16$$

$$E = 0.16$$

The table value of Z at 90% level is $Z(1/2)$ is 1.645
(from table)

$$\begin{cases} \leq 0.90 \\ \frac{\alpha}{2} = 0.05 \\ Z = 1.645 \end{cases}$$

from the table

$$ME \text{ of } \bar{x} = M\bar{x} = Z(1/2) \cdot \frac{\sigma}{\sqrt{n}}$$

$$= 1.645 \times \frac{0.16}{\sqrt{64}} = 0.329$$

90% confidence limits for $\mu = \left(\bar{x} - Z(1/2) \frac{\sigma}{\sqrt{n}} ; \bar{x} + Z(1/2) \frac{\sigma}{\sqrt{n}} \right)$
 $= 31.671 ; 32.329$

Q. A sample size of 300 was taken whose variance is 2254
mean is 50. Find the max. error to estimate the population mean and also construct 95% confidence interval for the mean.

n = 300

$$\sigma^2 = 22.5$$

$$\sigma = 15$$

$$Z(1/2) = 1.96$$

$$\bar{x} = 54$$

$$M.E \text{ or A.M.E} = Z(1/2) \cdot \frac{\sigma}{\sqrt{n}}$$

$$= 1.96 \times \frac{15}{\sqrt{300}} = 1.6974$$

95% confidence limits for μ : $(\bar{x} - Z(1/2) \cdot \frac{\sigma}{\sqrt{n}} ; \bar{x} + Z(1/2) \cdot \frac{\sigma}{\sqrt{n}})$

$$= (52.3026 ; 55.6974)$$

Q. It is decided to estimate the mean no. of hrs to continue until a certain computer will first require repair if it can be assumed that ~~it fails~~ is how larger the sample will be needed so that one will be able to assert 90% confidence the sample mean of almost 10 hrs.

Qd)

Given $\bar{x} = 4.8$

$$Z(1/2) = 1.645$$

Max error ± 10

$$M.E = Z(1/2) \cdot \frac{\sigma}{\sqrt{n}}$$

$$\sqrt{n} = \frac{Z(1/2) \cdot \sigma}{M.E} \Rightarrow n = \left(\frac{Z(1/2) \cdot \sigma}{M.E} \right)^2 = \left(\frac{1.645 \times 4.8}{10} \right)^2 = 62.3068 \approx 62$$

3. In a study of an automobile insurance a random sample of 80 body repairs cost had a mean of 672.36 rs & S.D. of

62.35 rs. If \bar{x} is used as a point estimate to the true avg repair cost with what confidence we can assert the max error do not exceed 10 rs

Sol
Given $n=80$

$$\sigma = 62.35$$

$$5. \bar{x} = 10 \text{ 672.36}$$

$$M.E \leq 10$$

$$M.E = 2(\zeta/2) \frac{\sigma}{\sqrt{n}}$$

$$2(\zeta/2) \cdot \sigma = \sqrt{n} \cdot M.E \Rightarrow 2(\zeta/2)^2 \frac{\sqrt{n} \times M.E}{\sigma} = \frac{\sqrt{80} \times 10}{62.35} = 1.645$$

$$2(\zeta/2) = 1.64$$

$$(\zeta/2) = 0.485 \text{ from table}$$

$$\zeta = 0.8502$$

∴ 85% confidence to estimate the population mean

Bob
1. A company wanted to introduce a new plan of work and a survey was conducted for this purpose out of 500 workers in 1 group

62% favoured the new plan and another group of sample of 400 workers 41% were against new plan construct 95% confidence limits for the difference of two proportions

Sol

$$\text{Given } n_1 = 500$$

$$n_2 = 400$$

p_1 = proportion of favoured new plan

$$= 0.62$$

P_2 = proportion of favoured of new plan in 2nd group

$$= 1 - 0.49 = 0.59$$

Population proportion $P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$

$$= \frac{500(0.62) + 400(0.59)}{900} = 0.6067$$

$$Q = 1 - P = 1 - 0.6067 = 0.3933$$

At 95% level $Z(\alpha/2) = 1.96$

95% confidence limits to difference of two proportions

$$\left[(P_1 - P_2) - Z(\alpha/2) \sqrt{PQ \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}, (P_1 - P_2) + Z(\alpha/2) \sqrt{PQ \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \right]$$

$$\left[(0.62 - 0.59) - 1.96 \sqrt{0.2386 \left(\frac{1}{500} + \frac{1}{400} \right)}, (0.62 - 0.59) + 1.96 \sqrt{0.2386 \left(\frac{1}{500} + \frac{1}{400} \right)} \right]$$

$$\left[-0.0342, 0.0942 \right]$$

2. A machine produced 20 defective articles in a batch of 400 after overhauling it produced 10 defectives in a batch of 300 construct 99% confidence limits for the difference of 2 proportions

99% confidence limits for the difference of 2 proportions

so

$$n_1 = 400$$

$$n_2 = 300$$

$$x_1 = 20$$

$$x_2 = 10$$

$$P_1 = \frac{x_1}{n_1} = 0.05$$

$$P_2 = \frac{x_2}{n_2} = 0.033$$

at 99% concentration $Z(\alpha/2)$ is 2.58

$$P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2} = \frac{400(0.05) + 300(0.033)}{700} = 0.0427$$

$$Q = 1 - P = 1 - 0.0427 = 0.9572$$

~~approx~~

$$\left[(P_1 - P_2) - z_{\alpha/2} \sqrt{PQ \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}, \quad \right]$$

$$\left[(0.05 - 0.033) - 2.58 \sqrt{0.0409 \left(\frac{1}{400} + \frac{1}{300} \right)}, \quad \right]$$

$$\left[-0.0232; \quad 0.0566 \right]$$

3. A sample of size 10 was taken from a population S.D of 0.03
find the max error with 99% confidence and also construct
99% limits for true avg with sample mean 20

sol

$$n = 10 \text{ (small sample)}$$

$$\bar{x} = 20$$

$$S.D = \sigma = 0.03$$

$$1 - \alpha = 99\% = 0.99$$

$$\alpha = 0.01$$

$$\frac{\alpha}{2} = 0.005$$

$$\left(z_{\alpha/2} \right) n - V = z_{0.005}^{(9)} \approx 1.833$$

$$M.E = z_{\alpha/2} \left[\frac{S}{\sqrt{n}} \right] = 0.0174$$

$$99\% \text{ confidence interval} = \left[\bar{x} - z_{\alpha/2} \left(\frac{S}{\sqrt{n}} \right); \bar{x} + z_{\alpha/2} \left(\frac{S}{\sqrt{n}} \right) \right]$$

$$\left[20 - 1.833 \left(\frac{0.03}{\sqrt{10}} \right) ; + \right]$$

$$(19.9326 ; 20.0174)$$

Ques

1. A random sample from a company very extensive files shows that the orders for a certain kind of missionary were filled with respectively in 10, 12, 19, 14, 15, 18, 11, 13 days. Use 1% level find m.e. errors and confidence limits for true avg.

Sol

$$n=8$$

$$\bar{x} = \frac{\sum x}{n} = \frac{10+12+19+14+15+18+11+13}{8} = 14$$

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

$$= \sqrt{\frac{(10-14)^2 + (12-14)^2 + (19-14)^2 + (14-14)^2 + (15-14)^2 + (18-14)^2 + (11-14)^2 + (13-14)^2}{7}}$$

$$= 3.2071$$

$$\alpha = 0.01, \frac{\alpha}{2} = 0.005$$

$$t_{(\alpha/2)}^{(n-1)} = t_{(0.005)}^{(7)} = 3.499 \text{ (From table)}$$

$$M.E = t_{(\alpha/2)} \cdot \frac{s}{\sqrt{n}} = 3.9674$$

$$1\%. \text{ Confidence limits} = \left[\bar{x} - t_{(\alpha/2)} \frac{s}{\sqrt{n}} ; \bar{x} + t_{(\alpha/2)} \frac{s}{\sqrt{n}} \right]$$

$$= 10.0326 ; 17.9674$$

2. A random sample of 10 students have the following IG.
 70, 120, 110, 101, 88, 83, 95, 98, 107, 100. Find the max error of
 5% level and also construct confidence interval for true avg

Sol

$$n = 10$$

$$\bar{x} = \frac{\sum x}{n} = 97.2$$

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

$$= \frac{(70-97.2)^2 + (120-97.2)^2 + (110-97.2)^2 + (101-97.2)^2 + (88-97.2)^2 + (83-97.2)^2 + (95-97.2)^2 + (98-97.2)^2 + (107-97.2)^2 + (100-97.2)^2}{9}$$

$$= 142.735$$

$$(43.8413 + 3.4258 = 17.3271)$$

$$\alpha = 0.05$$

$$\frac{\alpha}{2} = 0.025$$

$$t_{(1/2)} \cdot \frac{s}{\sqrt{n}} = 2.22$$

$$ME = t_{(1/2)} \cdot \frac{s}{\sqrt{n}} = 10.2099$$

$$5\% \text{ confidence limits} = (\bar{x} - t_{(1/2)} \cdot \frac{s}{\sqrt{n}} ; \bar{x} + t_{(1/2)} \cdot \frac{s}{\sqrt{n}})$$

$$= (86.9901 ; 107.4099)$$

3. Two independent samples of 8 and 7 items respectively have the following values. Find the confidence limits at 5% level for the diff. of two avg.

$$(i) \text{ Sample 1: } 11, 11, 13, 11, 15, 17, 12, 14 \quad \bar{x} = 12$$

$$(ii) \text{ Sample 2: } 9, 11, 10, 13, 9, 7, 10 \quad \bar{y} = 10$$

$$n_1=8, n_2=7$$

$$(x-\bar{x})^2: 1, 1, 4, 1, 9, 49, 0, 4 \rightarrow 26$$

$$(y-\bar{y})^2: 1, 1, 0, 9, 1, 4, 0 \rightarrow 16$$

$$S_d = \sqrt{\frac{\sum (x_i - \bar{x})^2 + \sum (y_i - \bar{y})^2}{n_1 + n_2 - 2}} = \sqrt{\frac{26 + 16}{8+7-2}} = 1.7974$$

$$\frac{(n_1+n_2-2)}{t_{\alpha/2}} = \frac{15}{2.160} = 7.0093$$

$$\text{C.I. (lower confidence limits)} = \left[(\bar{x} - \bar{y}) \pm t_{\alpha/2} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right] \\ = [12 - 10] \pm 7.0093 \sqrt{\frac{1}{8} + \frac{1}{7}} \\ = 2 \pm 7.0093 \times 0.4025 \\ = 2 \pm 2.8166 \\ = [-0.8166, 4.8166]$$