باب
3

ڈیٹا کی تنظیم کاری

اس باب کا مطالعہ کرنے کے بعد اتنی استعداد ہو جانی چاہیئے کہ آپ:

- مزید شماریاتی تجزیے کے لیے ڈیٹا کی درجہ بندی کرسکیں گے؛
- مقداری اور کیفیتی درجہ بندی کے درمیان امتیاز کرسکیں گے؛
- تواتری تقسیم کا جدول تیار کرسکیں گے؛
- کلاسوں کو وضع کرنے کی تکنیک جان سکیں گے؛
- ملان کے نشان زدگی طریقے (Tally marking method) جان سکیں گے۔
- یک متغیرہ (univariate) اور دو متغیرہ (bivariate) تواتری تقسیمی کے درمیان امتیاز کرسکیں گے۔

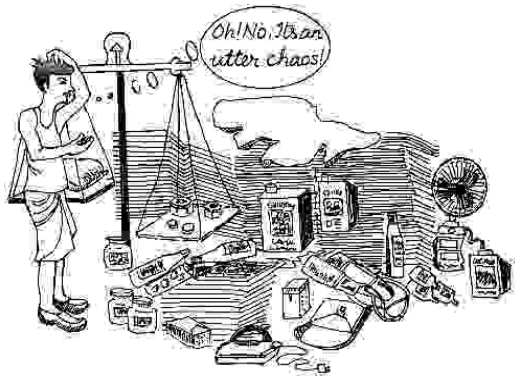
1. تعارف

پچھلے باب میں آپ نے سیکھا کہ ڈیٹا کس طرح جمع کیا جاتا ہے۔ آپ نے مردم شماری اور نمونہ کاری کے درمیان فرق بھی جانا۔ اس باب میں، آپ جانیں گے کہ ڈیٹا جو آپ نے جمع کیا ہے، اس کی درجہ بندی کیسے کی جاتی ہے۔ خام ڈیٹا (جس کا تجزیہ نہ کیا گیا ہو) کی درجہ بندی کا مقصد ان کو اس طرح منظّم کرنا ہے کہ مزید شماریاتی تجزیے میں آسانی ہو سکے۔

کیا آپ نے کبھی اپنے مقامی کباڑی والے جسے آپ پرانے اخبار، ٹوٹے پھوٹے گھریلو سامان، خالی بوتلیں، پلاسٹک وغیرہ فروخت کرتے ہیں، کا مشاہدہ کیا ہے۔ وہ ان چیزوں کو آپ

سے خریدتا ہے اور انھیں فروخت کر دیتا ہے جو ان کی باز گردانی (ان اشیا کو کارآمد بنانے کے لیے نئی شکل دینا) کرتے ہیں۔لیکن اس کی دوکان میں اتنے سارے کباڑے سے اس کے لیے اپنے کاروبار کو سنبھالنا مشکل ہوجاتا اگر وہ اسے مناسب طور پر ترتیب نہ دیتا۔اس صورتِ حال سے سہولت کے ساتھ نمٹنے کے لیے وہ مختلف کباڑوں کو مناسب طور پر گروپ بندی یا درجہ بندی کرتا ہے۔وہ پرانے اخباروں کو ایک ساتھ رکھتا ہے اور انھیں رسی سے باندھتا ہے۔ تمام خالی شیشے کی بوتلوں کو ایک بورے میں رکھتا ہے۔ وہ اپنی دوکان کے ایک کونے میں دھات کے سامان کا ڈھیر لگاتا ہے اور ''لوہا''، ''تانبہ''، ''المونیم''، ''پیتل'' وغیرہ جیسے گروپوں میں ان کو چھانٹتا ہے۔اس طرح وہ اپنے کباڑ کو مختلف درجوں ''اخبارات''، ''پلاسٹک''، ''شیشہ''، ''دھات'' میں گروپ بندی کرتا ہے، اور انہیں اس طرح مرتب کرتا ہے۔ جب اس کا کباڑ مرتب اور درجہ بند ہوجاتا ہے تب اس کے لیے آسان ہوجاتا ہے کہ وہ کسی خریدار کے طلب کرنے پر کسی مخصوص سامان کو دریافت کر سکے۔

اسی طرح جب آپ اپنی اسکولی کتابوں کو ایک خاص ترتیب میں لگاتے ہیں تو اسے برتنا آپ کے لیے آسان ہوجاتا ہے۔آپ انھیں مضامین کے لحاظ سے درجہ بند کر سکتے ہیں جہاں ہر مضمون ایک گروپ یا ایک درجہ بن جاتا ہے۔لہذا جب آپ کو تاریخ پر کسی مخصوص کتاب کی ضرورت ہوتی ہے بطور مثال آپ کو ''تاریخ'' گروپ میں رکھی کتاب کو تلاش کرنے کی ضرورت



پڑتی ہے۔ ورنہ آپ جس مخصوص کتاب کو تلاش کر رہے ہیں اس کا پتہ کرنے کے لیے پورے کتابوں کے ذخیرے کو تلاش کرنا ہوگا۔

جب ہم اشیاء یا چیزوں کی درجہ بندی کرتے ہیں تو ہمارا قیمتی وقت اور محنت بچتی ہے۔اسے کسی بے ترتیب طریقے سے نہیں انجام دیا جاتا۔ کباڑی والا اپنے کباڑ کا اس طرح گروپ بناتا ہے کہ ہر گروپ ایک جیسی مدوں پر مشتمل ہو۔مثال کے لیے شیشے کے گروپ کے تحت اسے خالی بوتلیں، ٹوٹے پھوٹے آئینے اور کھڑکی کے شیشے وغیرہ رکھنے ہوتے ہیں۔اسی طرح جب آپ اپنی تاریخ کی کتابوں کو ''تاریخ گروپ'' کے تحت درجہ بند کرتے ہیں تو آپ اس گروپ میں مختلف مضمون کی کتابیں نہیں رکھیں

سر گرمی

- اپنے مقامی ڈاکخانے یہ دریافت کرنے کے لیے جائیں کہ خطوط کو کس طرح چھانٹا جاتا ہے۔کیا آپ جانتے ہیں کہ کسی خط میں پن کوڈ کس بات کی علامت ہوتا ہے؟ اپنے ڈاکیے سے دریافت کریں۔

گے۔ ورنہ گروپ بنانے کا پورا مقصد بے کار ہو جائے گا۔ لہذا، درجہ بندی گروپوں یا درجوں میں یکساں چیزوں کو مرتب کرنے یا منظّم کرنے کو کہا جاتا ہے۔

2. خام ڈیٹا

کباڑے والے کے کباڑ کی طرح ہی غیر درجہ بند ڈیٹا یا خام ڈیٹا نہایت غیر منظّم ہوتے ہیں۔ یہ اکثر بہت طویل ہوتے ہیں اور انھیں برتنا باعث زحمت ہوتا ہے۔ ان سے بامعنی نتائج نکالنا بہت تھکا دینے والا کام ہوگا کیوں کہ یہ آسانی سے شماریاتی طریقے نہیں حاصل کرتے۔ اس لیے ایسے ڈیٹا کی مناسب تنظیم اور پیشکش کی ضرورت کسی نظامی شماریاتی تجزیے کو انجام دینے سے پہلے ہوتی ہے لہذا ڈیٹا کو جمع کرنے کے بعد اگلا قدم ایک درجہ بند شکل میں انھیں منظّم کرنا اور پیش کرنا ہے۔

مان لیجیے آپ ریاضیات میں طلبا کی کارکردگی جاننا چاہتے ہیں اور آپ نے اپنے اسکول کے 100 طلبا کے ریاضی میں حاصل کیے گئے نمبروں پر مبنی ڈیٹا جمع کیا ہے۔ اگر آپ اسے جدول کے طور پر پیش کرتے ہیں تو کچھ جدول 3.1 کی طرح دکھائی دے سکتا ہے۔

جدول 3.1

کسی امتحان میں 100 طلبا کے ذریعہ حاصل کیے گئے ریاضی کے نمبر

40	49	100	66	56	51	60	10	45	47
41	51	55	59	48	62	55	56	59	60
50	62	65	57	59	50	66	64	69	42
90	55	14	20	56	17	75	37	30	64
54	56	49	90	34	25	14	55	51	62
66	65	85	60	82	40	82	49	47	70
65	55	28	12	48	70	69	64	44	49
22	14	56	0	80	71	41	25	40	49
35	30	12	59	61	43	70	46	53	66
25	90	14	32	39	82	76	57	44	45

یا آپ اپنے پڑوس میں غذا پر ان کے اوسط اخراجات معلوم کرنے کے لیے 50 گھروں کی غذا پر ہونے والے ماہانہ اخراجات پر ڈیٹا اکٹھا کر سکتے ہیں۔ اس معاملے میں جمع کیے گئے ڈیٹا کو جدول کی شکل میں پیش کیا جا سکتا تھا جو کہ جدول 3.2 سے مشابہ ہوگا۔

جدول 3.2

50 گھروں کی غذا پر ماہانہ گھریلو اخراجات (روپے میں)

2760	1735	3473	1559	1904
4439	1855	1753	1612	2041
1523	2346	1823	1085	5090
1183	2152	1110	1360	1211
2712	2628	1105	1315	1218
1171	1183	1264	1812	4248
2048	1137	1953	1180	1007
3676	2621	1324	1583	2025
2575	2177	1962	1832	1397
1396	3222	1146	1365	1293

جدول 3.1 اور جدول 3.2 دونوں خام یا غیر درجہ ڈیٹا ہیں۔ دونوں جدول میں آپ پاتے ہیں کہ اعداد کسی ترتیب میں نہیں منظّم کیے گئے ہیں۔ اب اگر آپ سے پوچھا جائے کہ جدول 3.1 سے ریاضی میں حاصل کیا گیا سب سے زیادہ نمبر کیا ہے۔ تب آپ کو 100 طلبا کے حاصل نمبروں کو پہلے یا تو عروجی (Ascending) یا نزولی (Desending) ترتیب میں منظّم کرنا ہوگا۔ یعنی یہ اکتاہٹ والا کام ہوتا ہے۔ یہ زیادہ تھکا دینے والا کام بن جاتا ہے اگر آپ کو 100 کے بجائے 1000 طلبا کو برتنا ہو۔ اسی طرح جدول 3.2 میں آپ غور کریں گے کہ آپ کے لیے یہ تحقیق کرنا مشکل ثابت ہوگا کہ 50 گھروں کا اوسط ماہانہ اخراجات کیا ہیں۔ اگر یہ تعداد بڑی ہوتی بالفرض 5000 گھروں کی تو یہ مشکل کئی گنا بڑھ جاتی ہے بالکل ہمارے کباڑی والے کی طرح جسے کسی مخصوص مد کو دریافت کرنے کے لیے نڈھال ہو جانا پڑتا اگر اس کا کباڑ بہت زیادہ اور غیر مرتب ہوتا۔ آپ کو اسی طرح کی صورت حال کا سامنا کرنا پڑ سکتا ہے اگر آپ کسی خام ڈیٹا سے جو کہ کافی بڑا ہے، کوئی معلومات حاصل کرنے کی کوشش کریں۔ اگر ہم اسے ایک جملے میں کہیں تو یہ کہیں گے کہ کسی بڑے غیر درجہ بند ڈیٹا سے معلومات حاصل کرنے کا کام بہت تھکا دینے والا کام ثابت ہوگا۔ خام ڈیٹا کی تلخیص کی جاتی ہے اور اس کی درجہ بندی کرنے کے ذریعہ اسے جامع و مکمل بنایا جاتا ہے۔ جب ایک جیسی خصوصیات کے حقائق ایک جیسے درجے میں رکھے جاتے ہیں تو انھیں کوئی بھی آسانی سے پتہ لگا سکتا ہے، ان کا موازنہ کیا جا سکتا ہے اور بغیر کسی دقت کے نتائج اخذ کیے جا سکتے ہیں۔ آپ نے باب 2 میں مطالعہ کیا کہ حکومت ہند ہر دس سال پر آبادی کی مردم شماری کا اہتمام کرتی ہے۔ مردم شماری کے خام ڈیٹا اتنے زیادہ مقدار میں اور متفرق اجزاء پر مشتمل ہوتے ہیں کہ ان ڈیٹا سے کوئی بامعنی نتیجہ نکالنا تقریباً ناممکن کام ہوتا ہے لیکن جب مردم شماری ڈیٹا کی درجہ بندی جنسی، تعلیمی، ازدواجی حیثیت، پیشے وغیرہ کے لحاظ سے کی جاتی ہے تب ہندوستان کی آبادی کی ساخت اور اس کی فطرت کو آسانی سے سمجھا جا سکتا ہے۔

خام ڈیٹا متغیرات پر اور مشاہدات پر مشتمل ہوتا ہے۔ خام ڈیٹا کی ہر اکائی ایک مشاہدہ ہوتی ہے۔ جدول 3.1 میں مشاہدہ

متغیرہ (variable) کی مخصوص قدر ''ریاضی میں طالب علم کے مارکس'' کو ظاہر کرتا ہے۔ خام ڈیٹا طالب علم کے مارکس پر 100 مشاہدات پر مشتمل ہے کیوں کہ اس میں 100 طلبا ہیں، جدول 3.2 خود متغیرہ ''غذا پر گھروں کے ماہانہ اخراجات'' کی مخصوص قدر کو ظاہر کرتا ہے۔ اس میں موجود خام ڈیٹا گھروں کے ماہانہ اخراجات کی مخصوص قدر کو ظاہر کرتا ہے۔ اس میں موجود خام ڈیٹا گھروں میں غذا پر ہونے والے ماہانہ اخراجات پر 50 مشاہدات پر مشتمل ہے کیوں کہ اس میں 50 گھر شامل ہیں۔

سرگرمی

- ایک سال کے لیے اپنی فیملی کے کل ہفتہ وار اخراجات کے ڈیٹا جمع کیجیے اور اسے جدول میں مرتب کیجیے۔ دیکھیے کہ آپ کے پاس کتنے مشاہدات ہیں۔ ڈیٹا تو ماہانہ مرتب کیجیے اور دیکھیے کہ مشاہدات کی کتنی تعداد ہے۔

3. ڈیٹا کی درجہ بندی

ایک درجہ بندی کے گروپ یا درجات مختلف طور پر انجام دیے جاسکتے ہیں۔ اپنی کتابوں کو مضامین ''تاریخ، جغرافیہ، ریاضی، سائنس وغیرہ کے لحاظ سے درجہ بندی کرنے کے بجائے آپ انھیں ایک ابجدی ترتیب میں مصنف کے لحاظ سے بھی درجہ بند کرسکتے ہیں۔ یا آپ اشاعت کی سال کے مطابق بھی درجہ بند کرسکتے ہیں۔ ان کی درجہ بندی کرنے کا جو طریقہ آپ چاہتے ہیں وہ آپ کی ضرورت کے لحاظ سے ہوگا۔

اسی طرح خام ڈیٹا کو مختلف طریقوں میں درجہ بند کیا جاسکتا ہے جو زیر غور مقصد پر منحصر ہوگا۔ انھیں وقت کے لحاظ سے بھی درجہ بند کیا جاسکتا ہے۔ اس طرح کی درجہ بندی کو تاریخ وار درجہ بندی (Chronological Classification) کے طور پر جانا جاتا ہے۔ اس طرح کی درجہ بندی میں ڈیٹا کو یا تو عروجی یا نزولی ترتیب میں وقت جیسے سال، تین مہینوں، مہینوں، ہفتوں وغیرہ کے حوالے کے ساتھ درجہ بند کیا جاتا ہے۔ درج ذیل مثال یہ ظاہر کرتی ہے کہ ہندوستان کی آبادی کی درجہ بندی سال کے لحاظ سے کی گئی ہے۔ متغیرہ 'آبادی' ایک سلسلہ وقت ہے کیوں کہ یہ مختلف سالوں کے لیے قدروں کے ایک سلسلے کی تصویر کشی کرتا ہے۔

مثال 1

ہندوستان کی آبادی (کروڑوں میں)

سال	آبادی (کروڑوں میں)
1951	35.7
1961	43.8
1971	54.6
1981	68.4
1991	81.8
2001	102.7

مکانی درجہ بندی (Spatial Classification) میں ڈیٹا کو جغرافیائی وقوع جیسے ملکوں، ریاستوں، شہروں، اضلاع

وغیرہ کے حوالے سے درجہ بند کیا جاتا ہے۔ مثال 2 میں مختلف ملکوں کی پیداوار دکھائی گئی ہے۔

مثال 2

مختلف ملکوں کے لیے گیہوں کی پیداوار

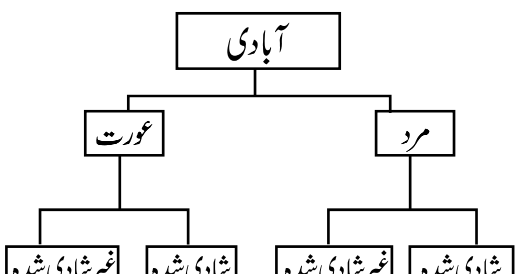
ملک	گیہوں کی پیداوار (کلوگرام/ایکڑ)
امریکہ	1925
برازیل	127
چین	893
ڈنمارک	225
فرانس	439
ہندوستان	862

سرگرمیاں

- مثال 1 سلسلہ وقت میں کون سے سال میں آپ ہندوستان کی آبادی سب سے کم پاتے ہیں۔ وہ سال بھی دریافت کیجیے جب اس کی آبادی سب سے زیادہ تھی۔
- مثال 2 میں وہ ملک دریافت کیجیے جس کی گیہوں کی پیداوار ہندوستان کی نسبت تھوڑی زیادہ ہے۔ فی صد کی اصطلاح میں یہ کتنا ہوگا؟
- پیداوار کی عروجی ترتیب میں مثال 2 کے ممالک مرتب کیجیے۔ کیا پیداوار کی نزولی ترتیب میں یہی مشق ہو سکتی ہے۔

کبھی کبھی آپ کو وہ خصوصیات حاصل ہوتی ہوں گی جنھیں مقداری طور پر نہیں ظاہر کیا جاسکتا ہے۔ اسی طرح کی خصوصیات کو کیفیتی یا توصیفی کہا جاتا ہے۔ مثال کے لیے قومیت، خواندگی، مذہب، جنس، ازدواجی حیثیت وغیرہ۔ تاہم ان اوصاف کی درجہ بندی کیفیتی خصوصیات کی موجودگی یا غیر موجودگی کی بنیاد پر کی جاسکتی ہے۔ اوصاف کی بنا پر ڈیٹا کی اس طرح کی درجہ بندی کو کیفیتی (Qualitative) درجہ بندی کہا جاتا ہے۔ درجہ ذیل مثال میں ہم پاتے ہیں کسی ملک کی گروپ بندی کیفیتی متغیرہ 'جنس' کی بنیاد پر کی گئی ہے۔ ایک مشاہدہ یا تو مرد ہوسکتا ہے یا عورت۔ ان دونوں خصوصیات کی مزید درجہ بندی زدواجی حیثیت (کیفیتی متغیرہ) کی بنیاد پر کی جاسکتی ہے۔ جیسا کہ ذیل میں دیا گیا ہے۔

مثال 3



پہلے مرحلے میں درجہ بندی وصف یعنی مرد یا غیر مرد (عورت) کی موجودگی اور غیر موجودگی پر مبنی ہوتی ہے۔ دوسرے مرحلے پر درجہ مرد اور عورت کی ذیلی تقسیم دوسرے وصف یعنی آیا وہ شادی شدہ ہے یا غیر شادی شدہ کی موجودگی یا غیر موجودگی کی بنیاد پر کی جاتی ہے۔

سر گرمی

- مثال 1 اور مثال 2 پر غور کریں اسے مقداری درجہ بندی کے طور پر بھی موزوں بنائیں اگرچہ ان کی درجہ بندی بالترتیب تاریخ وار اور مکانی طور پر کی گئی۔ یہ سوچنے کی کوشش کریں اور اسباب کا بھی اندازلگائیں کہ کیوں یہ مقداری درجہ بندیاں ہیں۔

جب کہ دوسری طرف لمبائی، وزن، عمر، آمدنی طلباء کے مارکس وغیرہ جیسی خصوصیات فطرتاً مقداری یا کمیتی ہیں ان خصوصیات کے ڈیٹا کو جب جمع کیا جاتا ہے تو ان کو درجوں میں گروپ بند کیا جاتا ہے اسے مقداری درجہ بندی (Quantitave Classification) کیا جاتا ہے۔

مثال 4

100 طلبا کے ریاضی میں حاصل نمبروں کی تواتری تقسیم

مارکس	تواتر
0-10	1
10-20	8
20-30	6
30-40	7
40-50	21
50-60	23
60-70	19
70-80	6
80-90	5
90-100	4
کل	100

مثال 4 جدول 3.1 میں دیے گئے 100 طلبا کے ریاضی میں مارکس کے ڈیٹا کی مقداری درجہ بندی تواتری تقسیم کے طور پر دکھائی گئی ہے۔

سر گرمی

- مثال 4 کے تواتر کی قدروں کو کل تواتر کے تناسب یا فی صد کے طور پر ظاہر کیجیے۔ نوٹ کیجیے کہ اس طریقے سے جو تواتر ظاہر کیا گیا ہے اسے نسبتی تواتر (Relation Frequency) کے طور پر جانا جاتا ہے۔
- مثال 4 میں کون سے درجے کو ڈیٹا کا سب سے زیادہ ارتکاز کیا گیا ہے؟ کل مشاہدات کے فی صد کے طور پر اسے ظاہر کریں۔ کون سے درجے پر سب سے کم ڈیٹا کا ارتکاز کیا گیا ہے؟

4. متغیرات: مسلسل اور مجرد

متغیرہ کی ایک آسان تعریف جو آپ نے پچھلے باب میں پڑھی تھی اس سے یہ نہیں پتہ چلتا کہ یہ کس طرح متنوع ہوتا ہے۔ مختلف متغیرات جداگانہ طور پر متغیر ہوتے ہیں جو اس بات پر منحصر ہے کہ ان کا تغیر کس طرح ہوتا ہے۔ انھیں موٹے طور پر دو قسم میں درجہ بند کیا جاتا ہے۔

1. مسلسل اور
2. مجرد یا غیر مسلسل

ایک مسلسل متغیرہ میں کوئی بھی ہندسی قدر لی جاسکتی ہے۔ اس میں سالم قدروں 1،2،3،4،...) کسری قدروں ($\frac{1}{2}$، $\frac{2}{3}$، $\frac{3}{4}$) اور ان قدروں کو لیا جاسکتا ہے جو کہ پوری طرح کسر (Fraction) نہیں ہیں جیسے $\sqrt{2} = 1.414$، [illegible]، $\sqrt{7} = 2.645$ ۔

مثال کے لیے طالب علم کا قد جیوں وہ بڑا ہوتا ہے بالفرض 90cm سے 150cm ہے، ان کے درمیان سبھی قدروں کو لیا جائے گا۔ یہ ان قدروں کو لے سکتا ہے جو صحیح اعداد جیسے 90cm، 100cm، 108cm، 150cm ہو۔ یہ کسری قدروں جیسے 90.85، 102.34،

104.99 وغیرہ جو کہ صحیح اعداد نہیں ہیں، انھیں اختیار کرسکتا ہے۔ اس طرح متغیرہ ''قد یا لمبائی'' ہر قرین قیاس قدر میں ظاہر ہونے کا اہل ہے اور اس کی قدروں کو قدروں کو لامحدود تدریج (gradation) میں تفصیلی تجزیہ بھی کیا جاسکتا ہے۔ مسلسل متغیرہ کی دیگر مثالیں وزن، وقت، دوری وغیرہ ہے۔

مسلسل متغیرہ کے برخلاف ایک مجرد یا غیر مسلسل متغیرہ صرف مخصوص قدریں اختیار کرسکتا ہے۔ اس کی قدر صرف محدود ''جست'' (یعنی یکبارگی) تبدیل ہوسکتی ہے۔ یہ ایک قدر سے دوسری قدر میں یکبارگی یا ایک دم تبدیل ہوجاتا ہے لیکن یہ ان کے درمیان کوئی وسطی قدر نہیں اختیار کرتا۔ مثال کے لیے مختلف کلاسوں میں ''کلاس میں طلبا کی تعداد'' جیسے متغیرہ ان اقدار کو اختیار کرے گا جو کہ صرف صحیح اعداد ہیں۔ یہ کوئی کسری قدر جیسے 0.5 نہیں اختیار کرسکتا کیوں کہ ''طالب علم کا نصف'' بے معنی ہے۔ لہذا یہ 25 اور 26 کے درمیان 25.5 جیسی قدر نہیں

اختیار کرسکتا بلکہ اس کے بجائے یا تو 25 یا 26 قدر ہوسکتی ہے۔ ہم جو مشاہدہ کرتے ہیں وہ یہ کہ چونکہ اس کی قدر 25 سے 26 میں تبدیل ہوتی ہے ان کے درمیان کی قدریں یعنی کسر میں اس کے ذریعہ نہیں اختیار کی جاسکیں۔ لیکن یہ تاثر نہیں ملتا کہ مجرد متغیرہ کسی کسری قدر کو نہیں اختیار کرسکتا۔ مان لیجیے x ایک متغیرہ ہے جو کہ $\frac{1}{8}$، $\frac{1}{16}$، $\frac{1}{32}$، $\frac{1}{64}$ ...، جیسی قدریں اختیار کرتا ہے تو کیا یہ ایک مجرد متغیرہ ہے؟ ہاں، کیوں کہ اگرچہ x کسری قدریں اختیار کرتا ہے اس لیے یہ دو متصل کسری قدروں کے درمیان کسی قدر کو نہیں اختیار کرسکتا۔ یہ $\frac{1}{8}$ سے $\frac{1}{16}$ میں اور $\frac{1}{16}$ سے $\frac{1}{32}$ میں ''ایک دم تبدیلی'' (Jumps) ہوتا ہے۔ لیکن $\frac{1}{8}$ اور $\frac{1}{16}$ یا $\frac{1}{16}$ اور $\frac{1}{32}$ کے درمیان قدر نہیں لے سکتا۔

سرگرمی

- مسلسل اور مجرد کے طور پر درج ذیل متغیرات میں امتیاز کیجیے
علاقہ، حجم، درجہ حرارت، پانسے (Dice) پر دکھائی دینے والا عدد، حاصل پیداوار، آبادی، بارش، سڑک پر کاروں کی تعداد، عمر۔

پہلے ہم نے ذکر کیا ہے کہ مثال 4، جدول 3.1 میں دکھائے گئے 100 طلبا کے ریاضی حاصل نمبروں کی تواتری تقسیم ہے۔ اس سے یہ ظاہر ہوتا ہے کہ کس طرح کلاسوں میں 100 طلبا کے 100 مارکس کی گروپ بندی کی گئی ہے۔ آپ کو حیرت ہوگی کہ کس طرح ہم نے جدول 3.1 کے خام ڈیٹا سے اسے حاصل کیا تھا لیکن اس سے پہلی کہ ہم اس سوال کا حل کریں، آپ کو یہ جاننا چاہیے کہ تواتری تقسیم کیا ہے۔

5. تواتری تقسیم کیا ہے

تواتری تقسیم (frequency distribution)، ایک مقداری متغیرہ کے خام ڈیٹا کو درجہ بند کرنے کا ایک جامع طریقہ ہے۔ یہ دکھاتی ہے کہ کس طرح متغیرہ یہاں طالب علم کے ذریعہ شمار کیے گئے ریاضی میں حاصل نمبر کو مختلف قدروں کو مختلف کلاسوں میں ان کے مطابق کلاس تواترات کے ساتھ تقسیم کیا جاتا ہے۔ اس معاملے میں ہمارے پاس مارکس کے دس درجات 0-10، 10-20، 100-90 ہیں۔ اصطلاح کلاس تواتر کا مطلب مخصوص کلاس میں قدروں کی تعداد ہے۔ مثال کے لیے کلاس 30-40 میں جدول 3.1 کے خام ڈیٹا سے مارکس کی 7 قدریں پاتے ہیں۔ یہ ہیں 30, 37, 34, 35, 39, 32 کلاس کا تواتر 30-40 اس طرح 72 ہے۔ لیکن آپ حیرت کریں گے کہ 40 کو جو خام ڈیٹا میں دوبارآ رہا ہے۔ 30 تا 40 کلاس میں نہیں شامل کیا گیا ہے۔ اگر اسے 30-40 کے تواتر میں شامل کیا جاتا تو 7 کے بجائے 9 آتا۔ اگر آپ اس باب کا مطالعہ بغور کریں گے اور تحمل اختیار کریں گے تو یہ پہیلی حل ہو جائے گی۔ آپ جواب خود ہی دریافت کرلیں گے۔

تواتری تقسیم جدول میں ہر کلاس، کلاس حدوں سے بندسی ہوتی ہے۔ کلاس کی حدیں کلاس کے دوسرے ہیں۔ کم ترین قدر کو کم تر کلاس حد کہا جاتا ہے اور سب سے زیادہ قدر کو اوپری کلاس حد کہا جاتا ہے۔ مثال کے لیے کلاس 70-60 کے لیے کلاس حدیں 60 اور 70 ہیں۔ اس کی نچلی کلاس حد 60 ہے اور اس کی اوپری کلاس حد 70 ہے۔ وقفہ کلاس یا وسعت کلاس اوپری کلاس حد اور نچلی کلاس حد کے درمیان فرق ہے۔ 70-60 کلاس کے لیے کلاس وقفہ 10 ہے۔ (اوپر کلاس حد نفی نچلی کلاس حد)

کلاس کا اوسط نقطہ یا کلاس مارک کلاس کی وسطی قدر ہے۔ یہ کلاس کی نچلی کلاس حد اور اوپری کلاس حد کے درمیان نصف فاصلے پر واقع ہوتا ہے اور درج ذیل طریقے سے اس کی تحقیق کی جاسکتی ہے۔

کلاس کا وسطی نقطہ یا کلاس مارک = (اوپری کلاس حد + نچلی کلاس حد / 2..................................(1)

کلاس مارک یا ہر کلاس کی وسطی قدر کا استعمال کلاس کی نمائندگی کے لیے کیا جاتا ہے جب ایک بار خام ڈیٹا کی کلاسوں (درجوں) میں گروپ کاری کی جاتی ہے، مزید شمار کے لیے انفرادی مشاہدات کا استعمال نہیں کیا جاتا۔ اس کے بجائے کلاس مارک کا استعمال کیا جاتا ہے۔

تواتری منحنی خط (frequency curve)، تواتری تقسیم

کی گرافی نمائندگی ہے۔شکل 3.1 ہماری درج بالا مثال میں ڈیٹا کی تواتری تقسیم کی ڈائیگرامی پیش کش دکھاتی ہے۔تواتری منحنی حاصل کرنے کے لیے ہم x- محور پر کلاس مارکس اور y- محور پر تواتر کا خاکہ کھینچتے ہیں۔

جدول 3.3

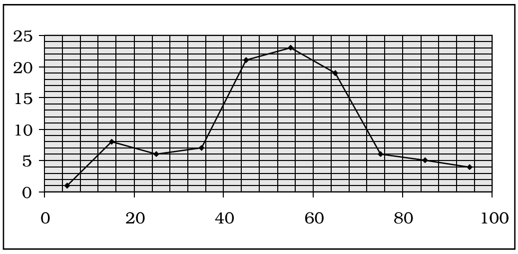
نچلی کلاس حدیں،اوپری کلاس حدیں اور کلاس مارک

کلاس	تواتر	نچلی کلاس حد	اوپری کلاس حد	کلاس مارکس
0-10	1	0	10	5
10-20	8	10	20	15
20-30	6	20	30	25
30-40	7	30	40	35
40-50	21	40	50	45
50-60	23	50	60	55
60-70	19	60	70	65
70-80	6	70	80	75
80-90	5	80	90	85
90-100	4	90	100	95

شکل 3.1 ڈیٹا کی تواتری تقسیم کی ڈائیگرامی پیش کش

تواتری تقسیم کیسے تیار کریں

جدول 3.1 کے خام ڈیٹا سے تواتری تقسیم تیار کرنے کے لیے درج ذیل سوالوں کو حل کیے جانے کی ضرورت ہے۔



1. کتنی کلاسیں (درجات) ہمارے پاس ہونے چاہئیں؟
2. ہر کلاس کی سائز کیا ہونی چاہیے؟
3. کلاس حدوں کو ہمیں کس طرح تواتر حاصل کرنا چاہیے۔
4. ہر کلاس کے لیے ہمیں کس طرح تواتر حاصل کرنا چاہیے۔

ہمارے پاس کتنی کلاسیں ہونی چاہئیں؟

کلاسوں کی تعداد متعین کرنے سے پہلے،ہمیں یہ معلوم کرنا چاہیے کہ ہمارے پاس موجود متغیرہ کسی حد تک قدر میں تبدیلیاں پیدا کرتا ہے۔متغیرہ کی قدر میں اس طرح کے تغیرات اس کی رینج کے ذریعہ اخذ کیے جاتے ہیں۔رینج متغیرہ کی سب سے بڑی اور سب سے چھوٹی قدروں کے درمیان فرق ہے۔ایک بڑی رینج اشارہ کرتی ہے کہ متغیرہ کی قدریں کافی وسیع ہیں۔جب کہ دوسری طرف ایک چھوٹی رینج ظاہر کرتی ہے متغیرہ کی قدروں کی توسیع کافی تنگ ہے۔ہماری مثال میں متغیرہ ''طالب علم کے مارکس'' کی رینج 100 ہے کیوں کہ کم ترین مارکس 0 اور بیش ترین مارکس 100 ہے۔اس سے یہ ظاہر ہوتا ہے کہ متغیرہ میں کافی تغیر ہے۔ رینج کی قدر حاصل کرنے کے بعد، جب ایک بار ہم وقفہ کلاس (Class in terval) کا فیصلہ کرتے ہیں تب کلاسوں کی تعداد کا

> **سرگرمیاں**
>
> درج ذیل کی رینج معلوم کیجیے
>
> - مثال 1 میں ہندوستان کی آبادی
> - مثال 2 میں گیہوں کی پیداوار

تعین کرنا آسان ہوجاتا ہے۔ نوٹ کیجیے کہ رینج سبھی کلاس کے وقفوں کی جمع ہے۔ اگر کلاس کے وقفے مساوی ہیں تب رینج کلاسوں کی تعداد کا حاصل ہے اور وقفہ کلاس ایک واحد کلاس ہے۔

رینج = کلاسوں کی تعداد × وقفہ کلاس.........2

رینج کی قدر دی گئی ہے، کلاسوں کی تعداد بڑی ہوگی اگر ہم چھوٹے کلاس وقفوں کا انتخاب کرتے ہیں۔ بہت زیادہ کلاسوں کے ساتھ تواتری تقسیم بھی بہت زیادہ متوقع ہوگی۔ ایسی تقسیم کو برتنا آسان نہیں ہے۔ لہٰذا ہمارے پاس ڈیٹا کا معقول دقیق مجموعہ ہونا چاہیئے۔ جب کہ رینج کی قدر دی گئی ہے، اگر ہم ایک ایسے وقفہ کلاس کا انتخاب کرتے ہیں جو کہ کافی بڑا ہے تب کلاسوں کی تعداد بہت کم ہوجاتی ہے۔ مجموعہ ڈیٹا تب بہت زیادہ دقیق ہوسکتا ہے اور ہم اس کے تنوع کے بارے میں معلومات کا نقصان نہیں اٹھانا چاہتے مثال کے لیے، بالفرض رینج 100 ہے اور وقفہ کلاس 50 ہے تب کلاسوں کی تعداد محض 02 (یعنی $\frac{100}{50} = 2$) ہوگی۔ حالانکہ کلاسوں کی تعداد کے تعین کے لیے کوئی متعین (بندھا ٹکا) اصول نہیں بس اس کے لیے اکثر عمومی ضابطے یا قیاس (rule of thumb) کا استعمال کیا جاتا ہے کہ کلاسوں کی تعداد 5 اور 15 کے درمیان ہونی چاہیے۔ ہماری مثال میں ہم نے 10 کلاسوں کو چنا ہے۔ چوں کہ رینج 100 ہے اور وقفہ کلاس 10 ہے اس لیے کلاسوں کی تعداد $\frac{100}{10} = 10$ ہے ہر کلاس کی سائز کیا ہوگی۔

اس سوال کا جواب پچھلے سوال کے جواب پر منحصر ہے۔ مساوات (2) دکھاتا ہے کہ متغیرہ کی رینج دی گئی ہے، جب ایک بار ہم وقفہ کلاس کا فیصلہ کرتے ہیں، تب کلاسوں کی تعداد کا تعین کرنا آسان ہوجاتا ہے۔ اسی طرح جب ایک بار ہم کلاسوں کی تعداد کا فیصلہ کرتے ہیں تب ہم وقفہ کلاس کا تعین کر سکتے ہیں۔ اس طرح ہم پاتے ہیں کہ یہ دو اخذ کردہ نتیجے ایک دوسرے کے ساتھ باہم مربوط ہیں۔ ہم کسی ایک کے بارے میں فیصلہ نہیں لے سکتے جب تک دوسرے پر نہ فیصلہ لیا جائے۔

مثال 4 میں ہمارے پاس کلاسوں کی تعداد 10 ہے۔ رینج کی قدر 100 دی گئی ہے، تب مساوات (2) کے حساب سے کلاس کے وقفے خود بخود 10 ہوجاتے ہیں۔ نوٹ کیجیے کہ موجودہ سیاق و سباق میں ہم نے جن کلاس وقفوں کو چنا ہے وہ قدر (حجم) میں مساوی ہیں۔ تاہم ہمیں ان کلاس وقفوں کو چننا پڑسکتا ہے جو کہ مساوی قدر کے نہیں ہوتے۔ ایسے معاملے میں کلاسوں کی غیر مساوی وسعت ہوتی ہے۔

کلاس حدوں کو ہمیں کس طرح متعین کرنا چاہیئے؟

جب ہم مسلسل متغیرہ کے خام ڈیٹا کی تواتری تقسیم کے طور پر درجہ بندی کرتے ہیں، ہم حقیقتاً کلاسوں میں انفرادی مشاہدات کی

گروپ کاری کرتے ہیں۔ کلاس کی اوپری کلاس حد کی قدر اس کلاس کی نچلی کلاس حد کی قدر کے ساتھ وقفہ کلاس شامل کرنے کے ذریعہ حاصل کرتے ہیں۔ مثال کے لیے کلاس 30-20 کی اوپری کلاس حد 10+20=30 ہے جہاں 20 نچلی کلاس حد ہے اور 10 وقفہ کلاس ہے۔ یہ طریقہ دیگر کلاسوں کے لیے دہرایا جاتا ہے۔

لیکن ہم پہلی کلاس کی نچلی کلاس حد کا فیصلہ کس طرح کرتے ہیں؟ یعنی کہ، 0 پہلی کلاس 10-0 کی نچلی کلاس کی حد کیوں ہے؟ پانسہ اس لیے ہے کہ ہم پہلی کلاس کی نچلی حد کے طور پر متغیرہ کی کم ترین قدر چنتے ہیں۔ درحقیقت ہمیں پہلی کلاس کی نچلی حد کے طور پر متغیرہ کی کم ترین قدر کی نسبت کم قدر چننا ہوتا ہے۔ اسی طرح آخری کلاس کی اوپری کلاس حد کے لیے ہمیں متغیرہ کی بیش ترین قدر کی نسبت زیادہ بڑی قدر چننا ہوگا۔ یہاں یہ غور کرنا اہم ہے کہ جب تواتری تقسیم کی تشکیل کی جاتی ہے تب کلاس کی حدوں کو اس طرح چننا چاہیئے کہ ہر کلاس کا وسطی نقطہ یا کلاس نشان جہاں تک ممکن ہو کوئی قدر جس کے قریب ڈیٹا مرتکز ہونے کی طرف مائل ہوتا ہو اس کے ساتھ انطباق (coincide) کرے۔

ہماری مثال میں 100 طلباء کے مارکس پر، ہم پہلی کلاس 10-0 کی نچلی حد کے طور پر 0 کو منتخب کرتے ہیں کیوں کہ کم ترین مارکس 0 تھا۔ اور یہی وجہ ہے کہ ہم اس کلاس کی نچلی کلاس حد کے عدد نہیں منتخب کر سکتے۔ اگر ہم نے ایسا کیا ہوتا تو ہم مشاہدہ 0 کو خارج کر دیتے تب پہلی کلاس 10-0 کی اوپری حد کلاس کی نچلی کلاس حد کے ساتھ وقفہ کلاس شامل کرنے کے ذریعہ حاصل کی جاتی۔ اس طرح پہلی کلاس کی اوپری کلاس 10=0-10 بن جاتی ہے اور اس طریق عمل کی پابندی دیگر کلاسوں کے لیے بھی کی جاتی ہے۔

کیا آپ نے کبھی غور کیا ہے کہ پہلی کلاس کی اوپر کلاس حد دوسری کلاس کی نچلی کلاس حد کے مساوی ہے؟ اور دونوں 10 کے مساوی ہیں۔ دیگر کلاسوں کے لیے اس کا مشاہدہ کیا جاتا ہے کیوں؟ اس کی وجہ یہ ہے کہ ہم نے خام ڈیٹا کی درجہ بندی کا اخراجی یا غیر شمولی (Exclusive Method) طریقہ استعمال کیا ہے۔ اس طریقے کے تحت ہم کلاسوں کی تشکیل اس طرح کرتے ہیں کہ ایک کلاس کی نچلی حد کا انطباق اس کی پچھلی کلاس کی اوپری کلاس حد کے ساتھ ہو۔

مسئلہ جس کا سامنا ہم آگے چل کر کریں گے وہ یہ ہے کہ ہم کسی مشاہدہ کی درجہ بندی کس طرح کرتے ہیں جو کہ کسی مخصوص کلاس کی اوپری کلاس حد کے نہ صرف مساوی ہو بلکہ اگلی کلاس کی نچلی کلاس حد کے بھی مساوی ہو۔ مثال کے لیے ہم پاتے ہیں کہ مشاہدہ 30 کلاس 30-20 کی اوپری کلاس حد کے مساوی ہے اور کلاس 40-30 کی نچلی کلاس حد کے بھی مساوی ہے۔ پھر دو کلاسوں 30-20 یا 40-30 یا کلاس 40-30 میں رکھ سکتے ہیں۔ یہ ایک ایسی صورت حال ہے (جس میں مسئلے کے دو متضاد حل ممکن ہوں) ہے جس میں کسی کو ہم پوش (O verlapping) کلاسوں میں ڈیٹا کی درجہ بندی کرتے وقت عام طور پر سامنا کرنا پڑتا ہے۔ اس مسئلے کو اخراجی یا غیر شمولی طریقے میں درجہ بندی

کے اصول کے ذریعہ حل کیا جاتا ہے۔

اخراجی یا غیر شمولی طریقہ

اس طریقے سے کلاسوں کی تشکیل اس طرح ہوتی ہے کہ ایک کلاس کی اوپری کلاس حد اگلی کلاس کی نچلی کلاس حد کے مساوی ہو۔ اس طریقے سے ڈیٹا کے تسلسل کو برقرار رکھا جاتا ہے۔ یہی وجہ ہے کہ درجہ بندی کا یہ طریقہ مسلسل متغیرہ کے ڈیٹا کے معاملے میں زیادہ موزوں ہے۔ اس طریقے کے تحت وقفے میں اوپری کلاس حد کو خارج کیا جاتا ہے لیکن کلاس کی نچلی حد کو شامل کیا جاتا ہے۔ اس طرح ایک مشاہدہ جو طریقے کے لحاظ سے اوپری کلاس حد کے بالکل مساوی ہو، وہ اس کلاس میں نہیں شامل ہوگا لیکن اگلی کلاس میں شامل کیا جائے گا۔ جب کہ دوسری طرح اگر یہ نچلی کلاس حد کے مساوی ہوتا تو اسے اس کلاس میں شامل کیا جاتا۔ طلبا کے مارکس کی ہماری مثال میں مشاہدہ 40 جو کہ جدول 3.1 کے خام ڈیٹا میں دوبار واقع ہوتا ہے، کلاس 40-30 میں نہیں شامل کیا گیا ہے۔ یہ اگلی کلاس 50-40 میں شامل کیا گیا ہے۔ یہی وجہ ہے کہ ہم پاتے ہیں کہ تواتر کلاس 40-30 کے مطابق 9 کے بجائے 7 ہے۔

کلاسوں کی تشکیل کرنے میں ایک اور طریقہ ہے اور اسے درجہ بندی کے شمولی طریقے (inclusive method) کے طور پر جانا جاتا ہے۔

شمولی طریقہ

غیر شمولی طریقے کے مقابلے میں شمولی طریقہ کلاس وقفے میں اوپری کلاس حد کو خارج نہیں کرتا۔ یہ کلاس میں اوپری کلاس کو شامل کرتا ہے۔ اس طرح دونوں کلاس حدیں کلاس وقفے کا حصہ ہیں۔

جدول 3.4

ایک کمپنی میں 550 ملازمین کی آمدنیوں کی تواتری تقسیم

آمدنی (روپے میں)	ملازمین کی تعداد
800-899	50
900-999	100
1000-1099	200
1100-1199	150
1200-1299	40
1300-1399	10
کل	550

مثال کے لیے، جدول 3.4 کی تواتری تقسیم میں کلاس 899-800 میں ان ملازمین کو شامل کرتے ہیں جن کی آمدنی یا تو 800 روپے ہے یا 800 اور 899 روپے کے درمیان یا 899 روپے ہے۔ اگر ملازم کی آمدنی 900 روپے ہی ہو تب اسے اگلی کلاس 999-900 میں رکھا جاتا ہے۔

کلاس وقفے میں تطابق

جدول 3.4 میں شمولی طریقے کا ایک قریبی مشاہدہ یہ دکھائے گا کہ اگرچہ متغیرہ ''آمدنی'' مسلسل متغیرہ ہے لیکن جب کلاسیں بنائی جاتی ہیں تو اس طرح کا تسلسل برقرار نہیں رہتا ہم کلاس کی اوپری حد اور اگلی کلاس کی نچلی حد کے درمیان فرق یا عدم تسلسل پاتے

ہیں۔مثال کے لیے پہلی کلاس 899 کی اوپری حد اور دوسری کلاس کی نچلی حد 900 کے درمیان ہم 1 کا فرق (gap) پاتے ہیں۔تب ہم ڈیٹا کی درجہ بندی کرتے وقت متغیرہ کے تسلسل کو کسی طرح یقینی بنا سکتے ہیں؟ یہ کلاس وقفے میں ایک تطابق (adjustment) کرنے کے ذریعہ حاصل کیا جاتا ہے۔ یہ تطابق درج ذیل طریقے سے انجام دیا جاتا ہے۔

1. دوسری کلاس کی نچلی حد اور پہلی کلاس کی اوپری حد کے درمیان فرق کا پتہ کریں۔مثال کے لیے جدول 3.4 میں دوسرے کلاس کی نچلی حد 900 ہے اور پہلی کلاس کی اوپری حد 899 ہے۔ ان کے درمیان 1 کا فرق ہے۔ یعنی (1=899-900)۔
2. (i) میں پائے گئی فرق کو دو سے تقسیم کریں یعنی ($\frac{1}{2} = 0.5$)۔
3. [2] میں حاصل کی قدر کو سبھی کلاسوں کی نچلی حدوں سے گھٹائیں (نچلی کلاس حد 0.5)
4. [2] میں حاصل کی گئی قدر کو سبھی کلاسوں کی اوپری حد میں جمع کریں (اوپری کلاس حد+0.5)

اس تطابق کے بعد جو کہ تواتری تقسیم میں ڈیٹا کا تسلسل بحال کرتا ہے، جدول 3.4 کو جدول 3.5 میں ترمیم کیا گیا ہے۔

جدول 3.5

ایک کمپنی کے 550 ملازمین کی آمدنیوں کی تواتری تقسیم

آمدنی (روپے میں)	ملازمین کی تعداد
799.5-899.5	50
899.5-999.5	100
999.5-1099.5	200
1099.5-1199.5	150
1199.5-1299.5	40
1299.5-1399.5	10
کل	550

کلاس حدوں میں تطابق کے ساتھ مساوات (1) جو کلاس مارک کا تعین کرتا ہے، اس کی ترمیم درج ذیل طور پر کی جائے گی۔

تطابق شدہ کلاس مارک = (تطابق شدہ اوپری کلاس حد تطابق شدہ نچلی کلاس حد)/2

ہمیں ہر کلاس کے تواتر کو کس طرح حاصل کرنا چاہیے

سادہ اصطلاح میں، مشاہدے کی تواتر کا مطلب ہے کہ خام ڈیٹا میں مشاہدہ کتنی بار واقع ہوتا ہے۔ ہمارے جدول 3.1 میں، ہم مشاہدہ کرتے ہیں کہ قدر 40 تین بار واقع ہوئی ہے، 0 اور 10 صرف ایک بار، 49 پانچ بار واقع ہوا ہے وغیرہ وغیرہ۔اس طرح 40 کا تواتر 3 ہے 0 کا 1، 10 کا 1 ہے، 49 کا 5 وغیرہ۔لیکن جب ڈیٹا کی گروپ بندی کلاسوں میں کی جاتی ہے جیسا کہ مثال 3 میں ہے ہے تب کلاس تواتر ایک مخصوص کلاس قدروں کی تعداد کی دلالت کرتا ہے۔مخصوص کلاس کے مقابل کلاس کے تواتر کا شمار ملان نشانوں کے ذریعہ انجام دیا جاتا ہے۔

ملان نشان لگانے کے ذریعہ کلاس تواتر دریافت کرنا

ہر طالب علم کے لیے جن کے مارکس اس کلاس میں شامل کیے گئے

جدول 3.6

ریاضی میں 100 طلبا کے مارکس کے ملان نشان کرنا

کلاس	مشاہدہ	ملان نشان (Tally Mark)	تواتر	کلاس مارک
0-10	0	I	1	5
10-20	10,14,17,12,14,12,14,14	𝍸 III	8	15
20-30	25,25,20,22,25,28	𝍸 I	6	25
30-40	30,37,34,39,32,30,35	𝍸 II	7	35
40-50	47,42,49,49,45,45,47,44,40,44,49,46, 41,40,43,48,48,49,49,40,41	𝍸 𝍸 𝍸 𝍸 I	21	45
50-60	59,51,53,56,55,57,55,51,50,56,59,56, 59,57,59,55,56,51,55,56,50,55,54	𝍸 𝍸 𝍸 III	23	55
60-70	60,64,62,66,69,64,64,60,66,69,62,61, 56,60,65,62,65,66,65	𝍸 𝍸 𝍸 IIII	19	65
70-80	70,75,70,76,70,71	𝍸 I	6	75
80-90	82,82,82,80,85	𝍸	5	85
90-100	90,100,90,90	IIII	4	95
	کل		100	

ہیں، ایک ملان نشان(tally)[/] کلاس کے مقابل لگایا جاتا ہے۔مثال کے لیے، اگر طالب علم کے حاصل کیے گئے مارکس 57 ہیں تب ہم ملان نشان[/] کو کلاس 50-60 کے مقابل لگاتے ہیں۔اگر مارکس 71 ہیں تو ملان نشان کلاس کو 70-80 کے مقابل لگایا جاتا ہے۔اگر کسی کو 40 مارکس ملے ہیں تو ملان کو کلاس 40-50 کے مقابل لگایا جاتا ہے۔جدول 3.6، جدول 3.1 سے ریاضی میں 100 طلبا کے مارکس کی ملان نشان زدگی دکھاتا ہے۔ملان کو شمار کرنا اس وقت آسان ہو جاتا ہے جب ان میں سے چار کو //// طور پر لگایا جاتا ہے اور پانچویں ملان نشان کو ان کے کاٹتے ہوئے 𝍸 اس طرح سے لگایا جاتا ہے۔ ملان نشانوں کو اس کے بعد پانچ کے گروپوں کے طور پر شمار کیا جاتا ہے۔لہذا، اگر کلاس میں 16 ملان نشان ہیں تب ہم آسانی کی خاطر اسے اس طرح لگاتے ہیں 𝍸، 𝍸، 𝍸، / اس طرح کلاس مع تواتر اس کلاس کے مقابل ملان نشانوں کی تعداد کے مساوی ہے

معلومات کا نقصان

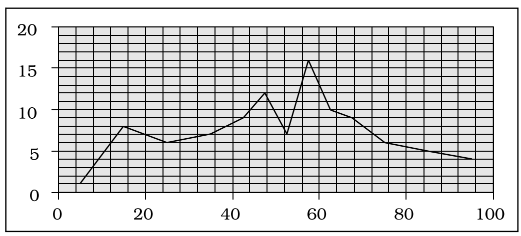
تواتری تقسیم کے طور پر ڈیٹا کی درجہ بندی میں خامیاں بھی ہوتی

ہیں د جب خام ڈیٹا کی تلخیص اسے جامع ومختصراور قابل فہم بناتے ہوئے کی جاتی ہے تب یہ خام ڈیٹا میں پائی جانے والی تفصیلات کو نہیں ظاہر کرتا۔ خام ڈیٹا کی درجہ بندی کرنے میں معلومات کا نقصان ہوتا ہے اگر چہ درجہ بند ڈیٹا کے طور پر اس کی تلخیص کر کے بہت کچھ حاصل کیا جاتا ہے۔ جب ایک بار ڈیٹا کو کلاسوں (درجوں) میں گروپ بند کرلیا جاتا ہے تب مزید شماریاتی حساب میں انفرادی مشاہدہ کی کوئی اہمیت نہیں ہوتی۔ مثال 4 میں کلاس 20-30، مشاہدات 25,22,20,25,25 اور 28 پر مشتمل ہے۔ لہذا جب ان ڈیٹا کو تواتری تقسیم میں ایک کلاس 20-30 کے طور پر گروپ بند کیا جاتا ہے تب موخرالذکر صرف کلاس (یعنی تواتر = 6) میں ریکارڈوں کی تعداد ان کی حقیقی قدریں فراہم کرتا ہے۔ اس کلاس میں سبھی قدروں کو کلاس وقفہ یا کلاس نشان (یعنی 25) کی وسطی قدر کے مساوی باور کیا جاتا ہے۔ مزید شماریاتی حساب صرف کلاس مارک کی قدروں پرمبنی ہے، نہ کہ اس کلاس میں مشاہدات کی قدروں پر۔ یہ بات دیگر کلاسوں کے لیے بھی صحیح ہے۔ اس طرح شماریاتی طریقوں میں مشاہدات کی حقیقی قدروں کے بجائے کلاس مارک کے استعمال سے معلومات کا نمایاں طور پر نقصان ہوتا ہے۔

غیر مساوی کلاسوں کے ساتھ تواتری تقسیم

اب تک آپ مساوی کلاس وقفوں کی تواتری تقسیم کے ساتھ واقف ہو چکے ہیں۔ آپ جانتے ہیں کہ کس طرح ان کی تشکیل خام ڈیٹا سے ہوتی ہے۔ لیکن بعض معاملوں میں تواتری تقسیم غیر مساوی کلاس وقفوں کے ساتھ ہوتی جو کہ زیادہ مناسب ہیں۔ اگر آپ مثال 4 کی تواتری تقسیم کا مشاہدہ کرتے ہیں جیسا کہ جدول 3.6 میں ہے تب آپ غور کریں گے کہ زیادہ تر مشاہدات کا ارتکاز 50-40، 60-50 اور 70-60 کلاسوں میں ہے۔ ان کے علی الترتیب تواتر 21، 23 اور 19 ہیں۔ اس کا مطلب یہ ہے کہ 100 مشاہدات میں سے 63 (19+23+21) مشاہدات ان کلاسوں میں مرتکز ہوئے ہیں۔ یہ کلاسیں مشاہدات کے ساتھ زیادہ گنجان واقع ہیں۔ اس طرح ڈیٹا کا 63 فی صد 40 اور 70 کے درمیان واقع ہیں۔ ڈیٹا کا باقی 37 فی صد کلاسوں 0-10، 10-20، 20-30، 30-40، 70-80، 80-90 اور 90-100 میں واقع ہیں۔ یہ کلاسیں مشاہدات کے ساتھ قلیل واقع ہیں۔ مزید آپ یہ بھی غور کریں گے کہ ان کلاسوں میں مشاہدات نسبتاً ان سے جو دیگر کلاسوں میں ہیں، اپنے متعلقہ کلاس مارکس زیادہ انحراف کرتے ہیں۔ لیکن اگر کلاسیں اس طرح تشکیل دی جاتی ہیں کہ کلاس کے مارکس جتنا ممکن ہوسکتا ہے کوئی قدر جس کے گرد کلاس میں موجود مشاہدات ارتکاز کی طرف مائل ہوتا ہے اس کے ساتھ انطباق کرے، تب اس صورت میں غیر مساوی کلاس وقفہ زیادہ مناسب ہے جدول 3.7 غیر مساوی کلاسوں کی اصطلاح میں جدول 3.6 کی وہی تقسیم ظاہر کرتا ہے۔ ہر ایک کلاس 40-50، 50-60 اور 60-70 دو کلاسوں میں تقسیم ہیں۔ کلاس 40-50، 40-45 اور 45-50 میں تقسیم

ہے، کلا50-60، 50-55 اور55-60 میں تقسیم ہے اور کلاس 60-70، 60-65 اور 65-70 میں تقسیم ہے۔ نئی کلاسوں 40-45، 45-50، 50-55، 55-60، 60-65، اور 65-70 کا کلاس وقفہ 5 ہے۔ دیگر کلاسوں0-10، 10-20، 20-30، 30-40، 70-80، 80-90 اور90-100 اپنے 10 کے کلاس وقفے کو برقرار رکھتے ہیں۔ اس جدول کا آخری کالم ان کلاسوں کے لیے کلاس مارکس کی نئی قدروں کو ظاہر کرتا ہے۔ جدول3.6 میں کلاس مارکس کی پرانی قدروں کے ساتھ ان کا



شکل3.2 تواتری منحنی خط

موازنہ کریں۔ غور کریں کہ ان کلاسوں میں مشاہدات اپنی نئی کلاس مارک قدروں کی نسبت اپنی پرانی کلاس مارک قدروں سے زیادہ انحراف کرتے ہیں۔ اس طرح نئی کلاس مارک قدریں پرانی قدروں کی نسبت ان کلاسوں میں ڈیٹا کی زیادہ نمائندگی کرتی ہیں۔

شکل3.2 میں جدول3.7 کی تقسیم کا تواتری منحنی خط دکھایا گیا ہے۔ جدول کے کلاس مارکس x- محور پر کھینچے گئے ہیں جب کہ تواترات کی ترسیم y- محور پر کی گئی ہے۔

سرگرمی

اگر آپ شکل3.2 کا موازنہ شکل3.1 سے کرتے ہیں، تو آپ کیا مشاہدہ کرتے ہیں؟ کیا آپ ان کے درمیان کوئی فرق پاتے ہیں؟ کیا آپ اس کی وضاحت کر سکتے ہیں؟

تواتری صف

اب تک آپ نے ریاضی میں 100 طلبا کے فی صد مارکس کا استعمال کرتے ہوئے مسلسل متغیرہ کے لیے ڈیٹا کی درجہ بندی پر بحث کی ہے۔ مجرد یا غیر مسلسل متغیرہ کے لیے اس کے ڈیٹا کی درجہ بندی کو تواتری صف کے طور پر جانا جاتا ہے۔ چوں کہ مجرد متغیرہ سالم قدروں کو اختیار کرتا ہے نہ کہ دو صحیح یا سالم قدروں کے درمیان وسطی کسری قدروں کو، اس لیے ہمارے پاس وہ تواترات ہوتے ہیں جو اس کی ہر سالم قدر کے مطابق ہوں۔

جدول3.8
خاندانوں کی جسامت کا تواتری صف

خاندانوں کی جسامت	خاندان کی تعداد
1	5
2	15
3	25
4	35
5	10
6	5
7	3
8	2
کل	100

جدول3.7

غیر مساوی کلاسوں کی تواتری تقسیم

کلاس	مشاہدہ	تواتر	کلاس مارک
0-10	0	1	5
10-20	10,14,17,12,14,12,14,14	8	15
20-30	25,25,20,22,25,28	6	25
30-40	30,37,34,39,32,30,35	7	35
40-45	42,44,40,44,41,40,43,40,41	9	42.5
45-50	47,49,49,45,45,47,49,46,48,48,49	12	47.5
50-55	51,53,51,50,51,50,54	7	52.5
55-60	59,56,55,57, 55, 56, 59,56, 59, 57, 59, 55,	16	57.5
	56,55,56,55		
60-65	60,64,62,64,64,60,62,61,60,62	10	62.5
65-70	66,69,66,69,66,65,65,66,65	9	67.5
70-80	70,75,70,76,70,71	6	75
80-90	82,82,82,80,85	5	85
90-100	90,100,90,90	4	95
کل		100	

جدول3.8 میں درج بالا مثال تواتری صف کو واضح کرتا ہے۔ متغیرہ خاندان کی جسامت ایک مجرد متغیرہ ہے جو صرف سالم قدروں کو اختیار کرتا ہے جیسا کہ جدول میں دکھایا گیا ہے۔ چوں کہ یہ دو متصل سالم قدروں کے درمیان کسی بھی کسری قدر کو نہیں اختیار کرتا اس لیے اس تواتری صف یہ کوئی کلاسیں یا درجات نہیں ہیں، یہاں کوئی کلاس وقفہ بھی نہیں ہوگا۔ چوں کہ مجرد تواتری تقسیم میں درجات

موجود نہیں ہوتے اس لیے کوئی کلاس مارک بھی نہیں ہوتا۔

جدول 3.9

20 فرموں کے فروخت کی دو تغیرہ تواتری تقسیم (لاکھ روپے میں) اور اشتہار کے اخراجات (ہزار روپوں میں)

	115-125	125-135	135-145	145-155	155-165	165-175	کل
62-64	2	1					3
64-66	1		3				4
66-68	1	1	2	1			5
68-70		2		2			4
70-72		1	1		1	1	4
کل	4	5	6	3	1	1	20

6. دو متغیرہ تواتری تقسیم

واحد متغیرہ کی تواتری تقسیم کو واحد یا یک متغیرہ تقسیم کہتے ہیں۔ مثال 3.3 واحد متغیرہ 'طالب علم کے مارکس' کی یک متغیرہ تقسیم کہا جاتا ہے۔ دو متغیرہ تواتری تقسیم دو متغیرات کی تواتری تقسیم ہے۔

جدول 3.9 دو متغیرات اور 20 کمپنیوں کے اشتہار کے اخراجات (لاکھ روپے میں) دکھاتا فروخت کی قدروں کو مختلف کالموں میں درجہ بندی کی گئی ہے اور اشتہار کے اخراجات کی قدروں مختلف قطاروں میں درجہ بند کیا گیا ہے۔ مہر خانہ متعلقہ قطار اور کالم قدروں کو دکھاتا ہے۔ مثال کے لیے 3 فرموں میں جن کے فروخت 135 تا 145 لاکھ ہیں اور ان کے اشتہار کے اخراجات 64 تا 66 ہزار ہیں۔ متغیرہ تقسیم کے استعمال کے بارے میں باب 8 ہم رشتگی (Correlation) میں بیان کریں گے۔

7. اختتام

ابتدائی اور ثانوی وسائل سے اکٹھا کیے گئے ڈیٹا خام یا غیر درجہ بند ہوتے ہیں۔ جب ایک بار ڈیٹا کو اکٹھا کیا جاتا ہے تب اگلا قدم مزید شماریاتی تجزیے کے لیے ان کی درجہ بندی کرتا ہے۔ درجہ بندی سے ڈیٹا میں تنظیم پیدا ہوتی ہے۔ اس باب کے ذریعہ آپ یہ جاننے کے اہل ہوجاتے ہیں کہ کسی طرح ڈیٹا کی درجہ بندی جامع انداز میں تواتری تقسیم کے ذریعہ کی جاسکتی ہے۔ جب آپ درجہ بندی کی تکنیکوں کے بارے میں واقف ہوجاتے ہیں تو آپ کے لیے مسلسل اور مجرد متغیرات دونوں کی تواتری تقسیم کی تخلیق کرنا آسان ہوجاتا ہے۔

خلاصہ

- خام ڈیٹا کی تنظیم درجہ بندی کے ذریعہ کی جاتی ہے۔
- تواتری تقسیم سے یہ ظاہر ہوتا ہے کہ کس طرح متغیرہ کی مختلف قدروں کومختلف کلاسوں میں ان کے موافق کلاس تواترات کے ساتھ تقسیم کیا جاتا ہے۔
- غیر شمولی یا اخراجی طریقے میں اوپری کلاس حد کو خارج کیا جاتا ہے لیکن نچلی کلاس حد کو شامل کیا جاتا ہے۔
- شمولی طریقے میں اوپری اور نچلی کلاس حدوں دونوں کو شامل کیا جاتا ہے۔
- تواتری تقسیم میں آگے کے شماریاتی حساب مشاہدات کی قدروں کے بجائے صرف کلاس مارک قدروں پر مبنی ہوتے ہیں۔
- کلاسوں (درجات) کی تشکیل اس طرح کی جاتی ہے کہ ہر کلاس کے کلاس مارک جتنا ممکن ہوسکتا ہے اس قدر کے قریب آتے ہیں جس کے گرد کلاس میں موجود مشاہدات مرتکز ہونے کی طرف مائل ہوتے ہیں۔

مشقیں

1. درج ذیل میں کون سے متبادلات صحیح ہیں۔

(i) کلاس کا وسطی نقطہ مساوی ہوتا ہے۔

(a) اوپری کلاس حد اور نچلی کلاس حد کی اوسط کے

(b) اوپر کلاس حد اور نچلی کلاس حد کے حاصل کے

(c) اوپری کلاس حد اور نچلی کلاس حد کے تناسب کے

(d) ان میں سے کوئی نہیں

(ii) دو متغیرات کی تواتر تقسیم کو جانا جاتا ہے۔

(a) یک متغیرہ تقسیم کے طور پر

(b) دو متغیرہ تقسیم کے طور پر

(c) کثیر متغیرات تقسیم کے طور پر

(d) ان میں سے کوئی نہیں

(iii) درجہ بند ڈیٹا میں شماریاتی حساب مبنی ہے

a) مشاہدات کی حقیقی قدروں پر

(b) اوپری کلاس حدوں پر

(c) نچلی کلاس حدوں پر

(d) کلاس کے وسطی نقاط پر

(iv) غیر شمولی یا اخراجی طریقے کے تحت

(a) کلاس کی اوپری کلاس حد کلاس وقفے میں خارج کی جاتی ہے

(b) کلاس وقفے میں کلاس کی اوپری کلاس حد کو شامل کیا جاتا ہے

(c) کلاس وقفے میں کلاس کی نچلی کلاس حد کو خارج کیا جاتا ہے

(d) کلاس وقفے میں کلاس کی نچلی کلاس حد کو شامل کیا جاتا ہے

(v) رینج ہے

(a) سب سے بڑے اور سب سے چھوٹے مشاہدات کے درمیان فرق

(b) سب سے چھوٹے اور سب سے بڑے مشاہدات کے درمیان فرق

(c) سب سے بڑے اور سب سے چھوٹے مشاہدات کی اوسط

(d) سب سے چھوٹے مشاہدے اور سب سے بڑے کی نسبت

2. چیزوں کی درجہ بندی میں کیا کوئی فائدہ ہے؟ اپنی روزمرہ کی زندگی سے کسی ایک مثال کے ساتھ وضاحت کیجیے۔

3. متغیرہ کیا ہے؟ مجرد اور مسلسل متغیرہ کے درمیان امتیاز کیجیے۔

4. ڈیٹا کی درجہ بندی میں استعمال کیے جانے والے غیر شمولی اور شمولی طریقوں کی توضیح کیجیے

5. جدول 3.2 میں ان ڈیٹا کا استعمال کیجیے جو 50 گھروں کی غذا پر ماہانہ گھریلو اخراجات (روپے میں) سے متعلق ہیں اور

(i) غذا پر ماہانہ گھریلو اخراجات کی رینج حاصل کیجیے

(ii) کلاس وقفوں کے موزوں تعداد میں رینج تقسیم کیجیے اور اخراجات کی تواتری تقسیم حاصل کیجیے

(vi) ان گھروں کی تعداد دریافت کیجیے جن کے غذا پر ماہانہ اخراجات ہیں۔

(a) 2000 روپے سے کم

(b) 3000 روپے سے کم

(c) 1500 روپے اور 2500 روپے کے درمیان

6. ایک شہر میں 45 فیملیوں کا سروے گھریلوں ساز و سامان جو وہ استعمال کرتے ہیں کی تعداد کے سروے کے لیے کیا گیا تھا۔ ان کے جواب کی بنا پر تواتری صف تیار کیجیے

3	3	3	3	2	2	1	2	1	2	2	2	2	3	1
3	5	1	5	1	2	6	1	6	2	2	3	2	3	3
3	4	1	3	0	2	4	3	4	2	4	7	2	4	2

جیسا کہ نیچے درج کیا گیا ہے

7. درجہ بند ڈیٹا میں معلومات کا نقصان کیا ہے؟

8. کیا آپ اس بات سے متفق ہیں کہ درجہ بند ڈیٹا خام ڈیٹا کے مقابلے بہتر ہے؟

9. یک متغیرہ اور دو متغیرہ تواتری تقسیم کے درمیان امتیاز کیجیے

10. درج ذیل ڈیٹا سے 7 کے کلاس وقفے لیتے ہوئے شمولی طریقے کے ذریعے تواتری تقسیم تیار کیجئے۔

6	4	9	2	7	12	18	27	23	21	29	22	15	17	28
9	15	21	27	33	4	8	12	16	20	5	10	3	8	1
	13	14	18	29	31	32	6	4	2	9	18	27	36	3
	25	19	20	24	26	28	32	37	5	1	7	9	11	15

مجوزہ سر گرمی

- اپنی پرانی مارک شیٹوں سے وہ مارکس دریافت کیجیے جو آپ نے پچھلی کلاسوں کے لیے ریاضی میں حاصل کیے ہیں۔ انھیں سال کے لحاظ سے مرتب کیجیے۔ جانچ کیجیے کہ آیا مارکس جو آپ نے مضمون میں حاصل کیا ہے وہ متغیرہ ہے یا نہیں۔ یہ بھی دیکھیے کہ ان سالوں میں آپ نے ریاضی میں اصلاح کی ہے۔